

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет Інформатики та обчислювальної техніки
Кафедра Автоматики та управління в технічних системах

«До захисту допущено»
Завідувач кафедру
_____ Ролік О. І.
(підпис) (ініціали, прізвище)
«__» _____ 2019 р.

Магістерська дисертація

зі спеціальності 126 Інформаційні системи та технології
на тему: «Підсистема розпізнавання об'єктів для автомобільних систем автономного керування»

Виконав: студент 6-го курсу, групи _____ ІА-82мп
(шифр групи)

_____ Хлівненко Михайло Вікторович
(прізвище, ім'я, по батькові) (підпис)

Науковий керівник _____ к.т.н, доцент каф. АУТС, Писаренко А. В.
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Консультант _____
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Рецензент _____
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Засвідчую, що у цій дипломній
роботі немає запозичень з праць
інших авторів без відповідних
посилань.

Студент _____
(підпис)

Київ – 2019 року

Національний технічний університет України
«Київський політехнічний інститут
імені Ігоря Сікорського»

Факультет _____ інформатики та обчислювальної техніки
(повна назва)

Кафедра _____ автоматики та управління в технічних системах
(повна назва)

Рівень вищої освіти – другий (магістерський) за освітньо-професійною
програмою

Спеціальність _____ 126 Інформаційні технології та системи
(код і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Ролік О. І.
(підпис) (ініціали, прізвище)

«__» _____ 2019 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Хлівненку Михайлу Вікторовичу

_____ (прізвище, ім'я, по батькові)

1. Тема дисертації «Підсистема розпізнавання об'єктів для автомобільних
систем автономного управління»

науковий керівник дисертації Писаренко Андрій Володимирович, к.т.н.,
(прізвище, ім'я, по-батькові, науковий ступінь, вчене звання)
доцент кафедри АУТС

затверджені наказом по університету від «__» _____ 2019 р. №__

2. Строк подання студентом дисертації _____

3. Об'єкт дослідження підсистема розпізнавання об'єктів для автомобільних
систем автономного керування з використанням швидкісних та ефективних
алгоритмів розпізнавання

4. Предмет дослідження ефективність, якість та швидкодія алгоритмів розпізнавання об'єктів.
5. Перелік завдання, які потрібно розробити огляд існуючих рішень, порівняння алгоритмів, модель системи, графічний матеріал
6. Орієнтовний перелік ілюстративного (графічного) матеріалу: схема електрична структурна, схема електрична функціональна, схема моделі, результати порівняння алгоритмів, приклади роботи алгоритмів та системи
7. Орієнтовний перелік публікацій: «Winter InfoCom Advanced Solutions 2019»,
8. Консультанти розділів дисертації
9. Дата видачі завдання – 02.09.2019

Календарний план

№ з/п	Назва етапів виконання дипломного проекту	Строк виконання етапів проекту	Примітка
1.	Огляд існуючих рішень	2.09.2019 р.	
2.	Опис та математична модель алгоритмів	14.09.2019 р.	
3.	Модель для порівняння алгоритмів	28.09.2019 р.	
4.	Аналіз результатів та вибір архітектури	1.10.2019 р.	
5.	Розроблення моделі системи	15.10.2019 р.	
6.	Аналіз отриманих результатів	31.10.2019 р.	
7.	Розроблення схеми структурної	5.11.2019 р.	
8.	Розроблення схеми функціональної	10.11.2019 р.	
9.	Розробка стартап – проекту	15.11.2019 р.	
10	Оформлення текстової документації	28.11.2019 р.	

Студент

Науковий керівник дисертації

<div style="text-align: right; margin-bottom: 5px;">Хлівненко М. В.</div> <div style="text-align: right; border-top: 1px solid black; margin-bottom: 5px;">(підпис)</div>	<div style="text-align: right; margin-bottom: 5px;">Писаренко А. В.</div> <div style="text-align: right; border-top: 1px solid black; margin-bottom: 5px;">(ініціали, прізвище)</div>
(підпис)	(ініціали, прізвище)

РЕФЕРАТ

Хлівненко М. В. Підсистема розпізнавання об'єктів для автомобільних автономного управління. КІП ім. Ігоря Сікорського, Київ, 2019.

Робота містить 131 с. тексту, 43 рисунків, 27 таблиць, 46 джерел та 3 додатки.

Враховуючи швидкий темп розвитку систем автономного керування, актуальним є дослідження показників якості навчання та роботи алгоритмів класифікації об'єктів та створення більш ефективних підходів.

Об'єктом розробки є підсистема розпізнавання об'єктів для автомобільних систем автономного керування з використанням швидкісних та ефективних алгоритмів розпізнавання та поєднанням методів отримання інформації для підвищення показників якості системи.

Метою магістерської дисертації є підвищення безпеки керування автомобілем, а також розвиток систем автономного керування за рахунок розроблення системи розпізнавання об'єктів з високими показниками якості.

Предметом дослідження є ефективність, якість та швидкодія алгоритмів розпізнавання об'єктів.

Ключові слова: системи автономного управління, розпізнавання об'єктів, YOLO, лідар, поєднання даних.

ABSTRACT

Khlivnenko M. An object recognition subsystem for automotive autonomous control systems. Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, 2019.

The work contains 131 p. of text, 43 figures, 27 tables, 46 references and 3 additions.

Taking into account the rapid pace of development of autonomous control systems, it is urgent to research the quality of training and the operation of algorithms for object classification and the creation of more effective approaches.

The object of development is the object recognition subsystem for automotive autonomous control systems that using high-speed and effective recognition algorithms and a combination of methods for obtaining information to improve the quality of the system.

The master's dissertation aims to improve the safety of driving, as well as the development of autonomous control systems by developing a system for recognition of objects with high quality indicators.

The subject of the research is the efficiency, quality and speed of object recognition algorithms.

Keywords: autonomous control systems, object detection, YOLO, lidar, data fusion.

ЗМІСТ

ПЕРЕЛІК ВИКОРИСТАНИХ СКОРОЧЕНЬ.....	9
ВСТУП.....	10
1 ПРИЗНАЧЕННЯ І ГАЛУЗЬ ЗАСТОСУВАННЯ.....	12
2 ОГЛЯД ІСНУЮЧИХ РІШЕНЬ	13
2.1 Загальна характеристика	13
2.2 Класифікація систем СОРО.....	15
2.2.1 Монохромні відеозображення.....	15
2.2.2 Кольорові зображення.	16
2.2.3 Інфрачервоні зображення	17
2.2.4 Комплексні системи розпізнавання об'єктів.....	18
2.3 Огляд алгоритмів.....	21
2.3.1 Алгоритми попередньої обробки.....	21
2.3.2 Алгоритми розпізнавання.....	23
2.4 Огляд існуючих систем.....	24
2.4.1 Pre Safe від Mercedes-Benz	24
2.4.2 Pre-collision system від Toyota	26
2.4.3 Pre-Sense від Audi	28
2.4.4 Driving Assistant від BMW.....	28
2.4.5 EyeSigh від Subaru	29
2.4.6 Tesla Autopilot	30
2.5 Дослідження ПНС.....	32
2.6 Висновки	35
3 ОПИС ТА МАТЕМАТИЧНА МОДЕЛЬ АЛГОРИТМІВ	36
3.1 Загальна характеристика	36
3.1.1 Регіони інтересу.....	37
3.1.2 Отримання ознак	38
3.1.3 Класифікація	42

3.2 Згорткові нейронні мережі	45
3.3 Двоетапні детектори	49
3.3.1 R-CNN.....	50
3.3.2 Fast R-CNN	53
3.3.3 Faster R-CNN.....	55
3.4 Одноетапні детектори	57
3.4.1 YOLO	58
3.5 Аналіз структури детекторів	59
3.6 Висновки	60
4 ПОРІВНЯННЯ АЛГОРИТМІВ	61
4.1 Опис макета.....	61
4.2 Отримані результати	63
4.3 Аналіз отриманих результатів	69
4.4 Висновки	69
5 КОНЦЕПЦІЯ СИСТЕМИ.....	70
5.1 Структура системи	70
5.2 Використаний алгоритм	71
5.3 Побудова об'ємних обмежувальних прямокутників	75
5.4 Функція втрат.....	75
5.5 Реалізація та отримані результати	76
5.6 Висновки	80
6 РОЗРОБЛЕННЯ СХЕМИ ЕЛЕКТРИЧНОЇ СТРУКТУРНОЇ	81
6.1 Опис	81
6.2 Висновки	83
7 РОЗРОБЛЕННЯ СХЕМИ ЕЛЕКТРИЧНОЇ ФУНКЦІОНАЛЬНОЇ	84
7.1 Опис	84
7.2 Висновки	87
8 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ.....	88
8.1 Опис ідеї проекту	88

8.2 Технологічний аудит ідеї проекту	92
8.3 Аналіз ринкових можливостей стартап-проекту	93
8.4 Розроблення ринкової стратегії проекту.....	101
8.5 Розроблення маркетингової програми стартап – проекту	104
8.6 Висновки	108
ВИСНОВКИ.....	110
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	112
Додаток А – Публікація	117
Додаток Б – Лістинг програми для аналізу алгоритмів	Ошибка! Закладка не определена.
Додаток В – Лістинг програми реалізації YOLO.	Ошибка! Закладка не определена.

ПЕРЕЛІК ВИКОРИСТАНИХ СКОРОЧЕНЬ

DL – Deep learning

FSD – Full self-driving

FPS – Frames per second

HOG – Histogram of oriented gradients

HW – Hardware

IIHS – Insurance institute for highway safety

IOU – Intersection over union

LDCF – Locally decorrelated channel features

R-CNN – Region based convolutional neural network

RGB – Red-green-blue

ROI – Region of interest

SVM – Support vector machine

RPN – Region proposal network

SSD – Single shot detector

YOLO – You only look once

СОРО – система оптичного розпізнавання об'єктів

ВСТУП

Останнім часом, зростає кількість дорожньо-транспортних пригод пов'язаних з пішоходами [1]. Факторами, що впливають на таку ситуацію є щільність міських забудов, значне збільшення кількості транспорту, нехтування правил дорожнього руху, неуважність водіїв та самих пішоходів тощо. Світові гіганти автомобілебудування приділяють багато уваги безпеці руху, кожного року представляють нові системи, вдосконалюють існуючі. Дані системи аналізують навколишній простір та у разі виникнення небезпечних ситуацій сповіщають про це водія, або самостійно втручаються у процес керування.

Крім того, з розвитком технологій все більшого розповсюдження зазнають системи автономного керування або, так звані, автопілоти. Такі системи здатні керувати автомобілем без втручання водія. Звичайно, щоб забезпечити високий рівень безпеки та комфорту, до таких систем висувають високі вимоги.

З огляду на такий стан речей, для розробки систем розпізнавання об'єктів використовують сучасні апаратні рішення, що здатні забезпечити усі необхідні вимоги. Проте, не менш важливою частиною систем є програмне забезпечення. Кожного року представляють нові алгоритми, методи та підходи до машинного навчання, кожен з яких, має свої переваги та недоліки. Проблема, що виникає перед розробниками – це необхідність вибору певного підходу та алгоритму, що забезпечив би при цьому усі висунуті вимоги, але за час створення, відлагоджування, тестування та сертифікації системи такі алгоритми зазвичай втрачають актуальність.

За останні роки проводилось мало порівняльних експериментів між різними алгоритмами, адже компаніям легше підтримувати існуючі системи просто модифікуючи їх.

Отже, постає питання порівняння сучасних методів розпізнавання та на основі аналізу їх показників створення такого підходу, який би був легко інтегрованим, при цьому мав високі показники якості роботи, мінімальний час навчання та мав можливість удосконалення у разі потреби.

Об'єктом розробки є підсистема розпізнавання об'єктів для автомобільних систем автономного керування з використанням швидкісних та ефективних алгоритмів розпізнавання.

Метою магістерської дисертації є підвищення безпеки кермування автомобілем за рахунок підсистеми розпізнавання об'єктів з високими показниками якості.

Предметом дослідження є якість алгоритмів розпізнавання об'єктів.

Для досягнення поставленої мети були сформульовані та вирішені наступні задачі:

- огляд та аналіз існуючих систем розпізнавання об'єктів;
- створення моделі для порівняння показників роботи різних алгоритмів розпізнавання;
- аналіз результатів дослідження;
- розроблення схеми електричної структурної;
- розроблення схеми електричної функціональної;
- представлення ефективного методу для розпізнавання об'єктів;
- розроблення програмної реалізації моделі системи та її аналіз.

Результати дослідження показників якості існуючих алгоритмів, а також опис запропонованого методу розпізнавання об'єктів були представлені на VIII міжнародній науково-практичній конференції «Winter InfoCom Advanced Solutions 2019», м. Київ, 2-3 грудня 2019 року. Додаток А.

Моделювання та аналіз показників якості алгоритмів здійснювалися за допомогою засобів пакету MATLAB. Розроблення та моделювання запропонованої системи здійснювалося з використанням мови програмування Python.

Магістерська дисертація складається з наступних розділів: вступ, призначення і галузь застосування, огляд стану справ, опис та математична модель алгоритмів, порівняння алгоритмів, модель системи, розроблення схеми електричної структурної, розроблення схеми електричної функціональної, розроблення стартап - проекту, висновки, список використаних джерел із 50 найменувань, 3 додатки.

Графічна частина включає 8 креслеників формату А3.

1 ПРИЗНАЧЕННЯ І ГАЛУЗЬ ЗАСТОСУВАННЯ

Розроблювана система призначена для використання у галузі автомобілебудування та створення систем автономного управління транспортними засобами. Головною метою системи є зменшення кількості дорожньо-транспортних пригод та підвищення загального рівня безпеки на дорогах, а також підвищення рівня комфорту управління автомобілем.

Головною задачею системи є швидке розпізнавання об'єктів, визначення їх розмірів та розрахунок положення у просторі. Для можливості використання у реальному житті необхідно забезпечити високий рівень надійності та якості роботи системи, а також забезпечити таку швидкість роботи, щоб система могла працювати у режимі реального часу.

Розроблювана система повинна використовувати уніфіковані та загальноприйняті стандарти обміну даними для реалізації можливості використання у існуючих транспортних засобах.

2 ОГЛЯД ІСНУЮЧИХ РІШЕНЬ

2.1 Загальна характеристика

При розробленні сучасних автомобілів велике значення приділяють системам безпеки. При цьому все більшу роль відіграють так звані інтелектуальні системи безпеки, які дозволяють гнучко реагувати на нестандартні ситуації. Існує цілий ряд інтелектуальних систем безпеки пасажирів, багато з яких вже знайшли своє застосування. Поряд з системами безпеки пасажирів велика увага приділяється розробці систем забезпечення безпеки пішоходів. Це нова галузь застосування інтелектуальних систем, що розвивається високими темпами.

Система розпізнавання пішоходів призначена для запобігання зіткнення з пішоходами. Система розпізнає людей біля автомобіля, автоматично уповільнює автомобіль, знижує силу удару і навіть уникає зіткнення. Застосування системи дозволяє на 20% зменшити смертність пішоходів при дорожньо-транспортній пригоді і на 30% знизити ризик важких травм [2].

Вперше система виявлення пішоходів була використана на автомобілях Volvo в 2010 році [3].

На даний момент існує кілька підходів до вирішення завдання забезпечення безпеки пішоходів, наприклад, встановлення спеціальних подушок безпеки на бампері автомобіля, примусова зупинка або зменшення швидкості автомобіля з метою запобігання аварії або подача попереджувального сигналу водієві. Проте ключовою проблемою перерахованих підходів є виявлення перешкоди і її ідентифікація.

В даному випадку – це виявлення об'єктів та місця їх розташування по ходу руху автомобіля. Серед завдань розпізнавання, проблема розпізнавання об'єктів є однією з найбільш складних.

Основні труднощі, пов'язані з вирішенням даного завдання, полягають в наступному: по-перше, фон, на якому необхідно шукати об'єкти, постійно змінюється, по-друге, положення об'єктів і відстань до них заздалегідь невідомі, по-третє, якість розпізнавання не повинна залежати від характеристик об'єкта.

З систем розпізнавання об'єктів на особливу увагу заслуговують системи оптичного розпізнавання об'єктів (СОРО). На даний момент існує більше 10 розробок в цій галузі. Склад і методи систем дуже різноманітні, при цьому, кожна з розробок має свої переваги і недоліки. Їх аналіз показує, що завдання виявлення об'єктів за оптичними даними знаходиться на самому ранньому етапі розвитку і представляє великий інтерес як з точки зору отримання нових наукових результатів, так і з точки зору практичного застосування.

Одними з основних характеристик СОРО є розміри робочої зони, які залежать від технічних характеристик оптичного пристрою. Як правило, глибина робочої зони не перевищує 30 метрів, а кут огляду 60° . Наскільки можна судити з опублікованих робіт, більшість СОРО мають схожі характеристики: відносно невелика глибина робочої зони легко з'ясовна: адже лінійні розміри об'єктів обернено пропорційні відстані до цих об'єктів. Такий великий розкид розмірів сильно ускладнює виявлення і розпізнавання об'єктів.

Другою важливою характеристикою СОРО є швидкодія. Очевидно, що система розпізнавання повинна працювати в реальному часі, отже, час обробки кадру має бути якомога меншим (зауважимо, що при швидкості 100 км / год машина проїжджає робочу зону СОРО менш ніж за одну секунду). Існують два підходи до вирішення даної проблеми: перший полягає в використанні спрощених, а отже, і більш швидких алгоритмів обробки зображень, що, звичайно ж, має відобразитися на якості.

Другий підхід полягає в тому, щоб використання складних алгоритмів роботи компенсувалося застосуванням спеціалізованих апаратних засобів, що дозволяють проводити обробку вхідної інформації за прийнятний час.

Звичайно, другий підхід вимагає великих фінансових витрат, але є і найбільш продуктивним. Ще однією характеристикою СОРО є діапазон освітленості об'єктів, в якому вона може працювати. В ідеальному випадку система розпізнавання повинна працювати як в джерелі яскравого світла (яскраве сонце), так і в умовах недостатньої освітленості.

Для подолання даної проблеми в різних дослідженнях пропонується використовувати або інфрачервоні камери, або різного роду радари.

2.2 Класифікація систем СОРО

Класифікацію систем СОРО можна провести за типом використовуваних технічних засобів (Відеокамер). При такому розподілі отримуємо три типи вхідних зображень:

2.2.1 Монохромні відеозображення

Розробники даних систем навмисно ставлять себе в дуже жорсткі рамки, тому не варто від даних систем чекати найкращих результатів. Однак такі системи мають свої переваги. Використання в якості вхідної інформації чорно-білих зображень дозволяє застосовувати більш прості і швидкі алгоритми розпізнавання. До того ж для такого типу зображень існує цілий ряд стандартних алгоритмів, які можуть бути легко адаптовані для даного завдання. Не останню роль відіграє також дешевизна відповідних апаратних засобів, тому в СОРО даного типу не застосовують великих розширень і, як правило, використовують зображення стандартного розміру. На рисунку 2.1 показано приклад роботи системи, що використовує монохромні зображення [4].

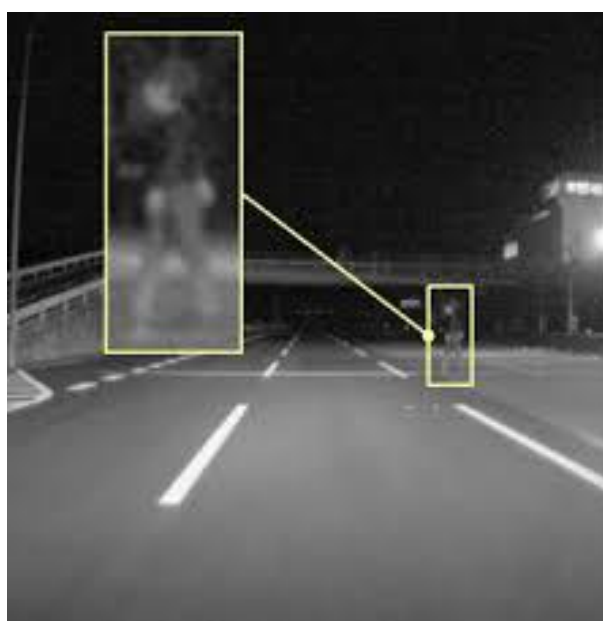


Рисунок 2.1 – Робота системи, що використовує монохромне зображення [4]

2.2.2 Кольорові зображення.

У порівнянні з чорно-білими зображеннями, кольорові несуть більше інформації про об'єкти і повинні істотно полегшувати їх виділення. У той же час з кольорового зображення завжди можна легко отримати чорно-біле. Тому для кольорових зображень істотно збільшується спектр можливих алгоритмів обробки від складних (квантування палітри, сегментація зображення та ін.) до простих алгоритмів обробки яскравості зображення. Тут необхідно зазначити, що складні алгоритми значно збільшують час розпізнавання, але більш точно виділяють об'єкт.

Для того щоб використовувати інформацію, яку несе в собі колір зображення, необхідно вміти порівнювати кольори. Дане завдання не має точного рішення. Існує безліч алгоритмів, але всі вони носять, здебільшого, емпіричний характер, і далеко не завжди застосовуються в даній області. Успішне вирішення завдання порівняння кольорів є ключовим при використанні кольорових зображень.

Деякі проблеми виникають при недостатньому освітленні, оскільки відеокамери в цих умовах не здатні добре передавати гамму кольорів. Тому на систему необхідно накладати додаткові обмеження по яскравості об'єкта дослідження (наприклад, використання даної системи тільки в світлий час доби), або ж в рамках системи розробити підсистему, яка буде працювати з чорно-білими зображеннями, що частково компенсує цей недолік.

Також, виникають проблеми при несприятливих погодних умовах, таких як дощ, снігопад та туман. В таких випадках необхідності забезпечити захист камери, адже краплі дощу створюють зайвий шум на знімках та можуть спотворювати зображення [5]. На рисунку 2.2 показано приклад роботи такої системи.

Аби підвищити точність розпізнавання при несприятливих умовах – використовують допоміжні засоби про які йтиметься далі.

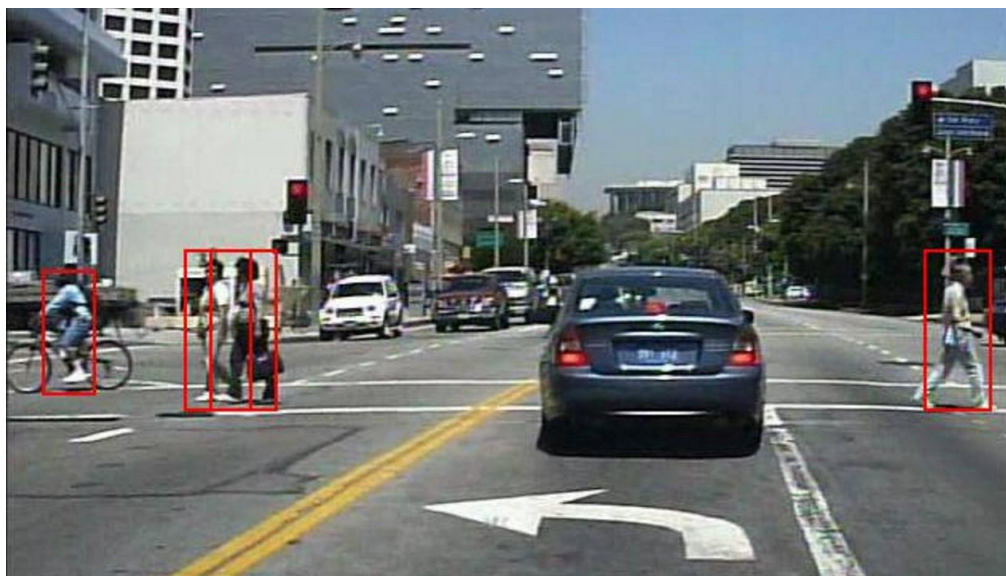


Рисунок 2.2 – Приклад роботи системи, що використовує кольорові зображення [5]

2.2.3 Інфрачервоні зображення

Перевагою систем, що працюють в інфрачервоному діапазоні, є можливість роботи в нічний час доби при мінімальному освітленні. Однак не зовсім зрозуміла стійкість таких систем при яскравому сонячному світлі і підвищеній температурі. Дослідження, показали, що при використанні зображень з малим розширенням, наприклад, 20x20, перевагу мають зображення, отримані в інфрачервоному діапазоні, водночас, при великому розширенні розпізнавання об'єктів краще проводити по чорно-білим зображенням. Однак, такий тип систем є досить популярним і зазвичай використовується у комплексі з іншими типами.

До переваг такого способу можна віднести те, що теплове випромінювання живого об'єкту можна побачити на відстані до 300-х метрів, що значно збільшує час на прийняття рішень, а також, теплове випромінювання можна побачити навіть за незначними перешкодами. Такий підхід зазвичай використовують в автомобільних системах нічного бачення та роботі в умовах недостатньої освітленості. Проте, такий підхід має ряд недоліків, серед яких є те, що такий тип камер не дозволяє визначати

неживі об'єкти, наприклад дерево, камінь тощо. На рисунку 2.3 показано приклад роботи системи, що використовує тепловізор [6].



Рисунок 2.3 – Приклад роботи системи з тепловізором [6]

2.2.4 Комплексні системи розпізнавання об'єктів

Існують системи, які для розпізнавання об'єкта використовують не тільки саме зображення від відеокамери, а й додаткову інформацію, що отримується за допомогою додаткових технічних засобів. Серед комплексних систем найцікавішими представляються наступні.

2.2.4.1 СОРО, що використовують стереопари

У таких системах використовуються дві відеокамери замість однієї. Стереопара використовується в основному для вимірювання відстані до об'єкта. Ця інформація значно полегшує виділення об'єктів на близькій відстані і відділення їх від фону. Встановлення стереопари легко реалізується практично. В таких СОРО, як правило, не застосовуються лінзи з сильною дисперсією, тому калібрування стереопари теж не представляє особливої проблеми. Але зате ціна такої системи зростає в кілька разів, оскільки обидві відеокамери, що використовуються в системі, повинні мати хорошу

роздільну здатність, що необхідно для якісного визначення відстані. Вимога для гарної роздільної здатності не випадкова, так як при малій роздільній здатності сильно ускладнюється завдання суміщення об'єктів на двох зображеннях. Виникають також проблеми з синхронізацією роботи відеокамер, але вони мають чисто технічний характер і легко вирішуються.

2.2.4.2 СОРО, що використовують радар

Призначення радару теж саме, що і стереопари. Тому в системах можна використовувати або радар, або стереопару, виходячи з наявних технічних можливостей. Використання радару спрощує алгоритми визначення відстані до об'єкта. Але його використання вимагає додаткових витрат, і дані, які знімаються з його допомогою, мають дещо інший формат, тому для їх обробки треба застосовувати додаткові алгоритми.

Але в роботах, що описують використання радару, не розкриваються технічні подробиці його застосування, такі як дальність, кути робочої зони і т. д.

Виходячи із розглянутих робіт, важко сказати, який з цих підходів є більш перспективним для вирішення задачі розпізнавання пішохода. Судячи з усього, вони обидва досить добре виконують завдання визначення відстані до об'єкта.

Перевагою стереопари є додаткова інформація про об'єкт, що з'являється за рахунок того, що камери сприймають об'єкт з різних точок. З іншого боку, на роботу радарів набагато менший вплив мають такі фактори, як освітленість, туман і т. п.

2.2.4.3 СОРО, що використовують набір послідовних знімків

Оскільки завдання стеження за об'єктом легше, ніж завдання знаходження об'єкта, то багато систем використовують результати своєї роботи на попередньому кроці для уточнення результатів на поточному. Досить очевидна стратегія, яка не вимагає ніяких додаткових апаратних витрат і яка досить просто реалізується. При використанні двох послідовних знімків також спрощується задача знаходження

рухомих об'єктів. В деяких випадках, при русі автомобіля, використовуючи два послідовні знімки, можна відокремити об'єкти, що знаходяться поблизу автомобіля, від віддалених об'єктів.

2.2.4.4 СОРО, що використовують лідар

В останні роки найпопулярнішим датчиком, що використовують для систем оптичного розпізнавання пішоходів є лідар.

Лідар надає можливість отримувати тривимірну модель оточуючого середовища, що має унікальні переваги для виявлення та відстеження виявленої перешкоди або об'єкту, вимірювання швидкості, навігації та позиціонування транспортного засобу. Динамічне виявлення та відстеження об'єктів дає змогу використовувати лідари не тільки для систем запобігання зіткнення, але і для інших систем автономного керування [7].

Однак лідар використовує лише числові значення, щоб робити висновок про тип об'єкта, що може призвести до низької ефективності, коли дані хмар точок є одиничними, або область виявлення заблокована. У деяких роботах [7] використовують алгоритм кластеризації транспортних засобів на основі відстані для ідентифікації транспортних засобів, що базуються на синтезі інформації, що містить дані про різні характеристики об'єкту, та використовують детермінований метод виконання цільової кореляції. Однак багатофункціональний синтез інформації досить громіздкий, правила не зрозумілі, і методи кореляції не можуть впоратися при появі нової цілі або зникненням старої. Через це, було запропоновано метод представлення на основі 3D-вокселів та використання методу дискримінаційного аналізу для моделювання перешкод. Цей метод є відносно новим і може забезпечити більш надійне виявлення статичних та рухомих об'єктів.

Лідар не може повною мірою надавати точну інформацію при туманній та спекотній погоді. Таким чином, продуктивність та точність розпізнавання виключно за допомогою лідара досить низька, а також необхідно враховувати складні дорожні

обставини у великих містах, все це не може гарантувати високі показники для автомобільних систем автономного керування.

Для адаптації до складності та мінливості дорожнього середовища використовують поєднання кольорової камери та лідару, щоб одночасно виявити ціль на автономному транспортному засобі, а потім надати достатню кількість інформації про положення об'єкта та інші його характеристики [8]. Приклад трьохвимірної моделі, отриманої за допомогою лідара показано на рисунку 2.4.

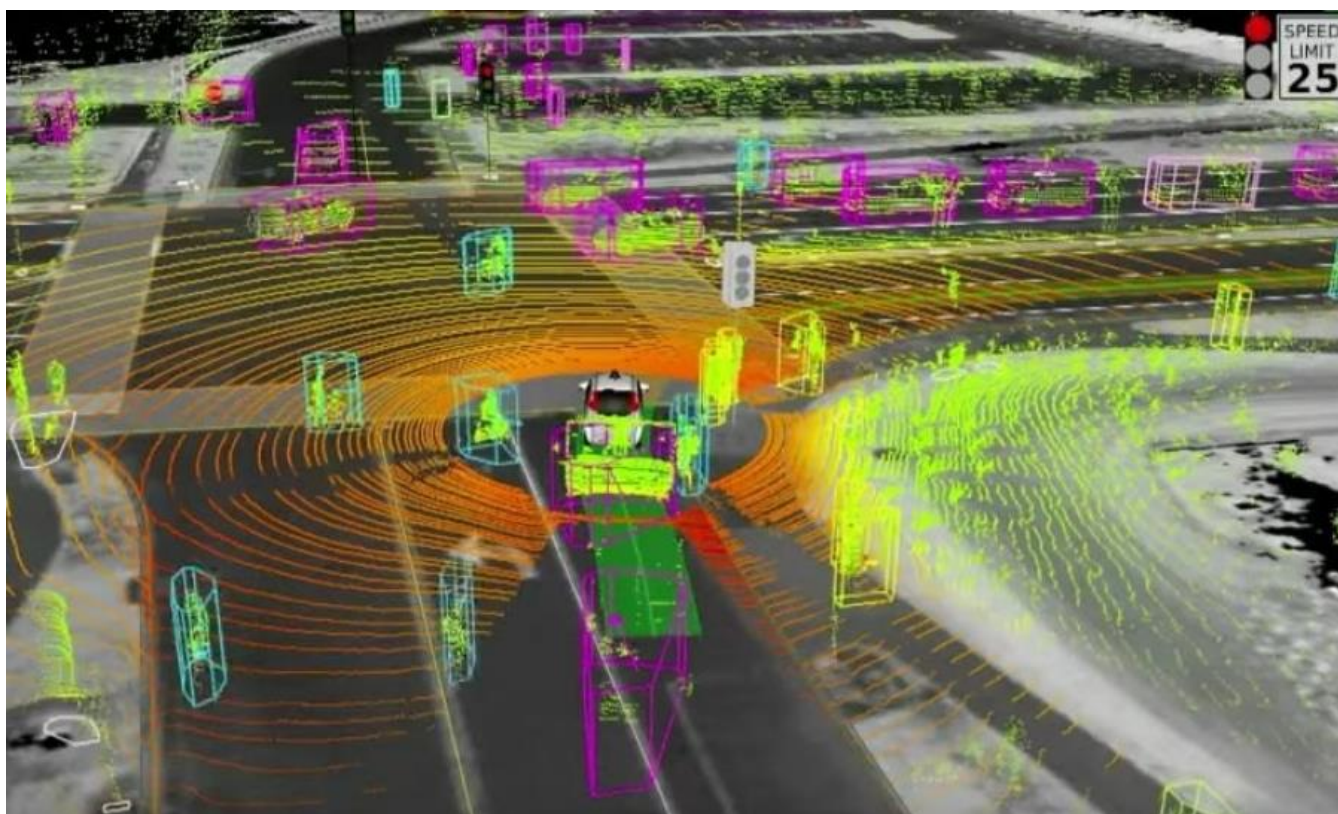


Рисунок 2.4 – 3D модель, отримана за допомогою лідара [8]

2.3 Огляд алгоритмів

2.3.1 Алгоритми попередньої обробки

Алгоритми попередньої обробки – найцікавіша і найрізноманітніша частина всіх СОРО. Головний метою в багатьох роботах є саме попередня обробка зображень. Набір використовуваних алгоритмів досить широкий. Єдиний етап, який присутній у

всіх проектах, це визначення контуру, хоча алгоритми отримання контурів в кожному проекті різні.

Якщо порівнювати алгоритми попередньої обробки і алгоритми розпізнавання, то найбільш слушну точку зору, висловлюють розробники Daimler Chrysler, які стверджують, що після отримання вектору ознак можна користуватися будь-яким із стандартних класифікаторів [9].

У кожній СОРО використовується цілий спектр алгоритмів обробки зображень, наприклад, в СОРО однієї з провідних фірм MobileEye [10] алгоритм попередньої обробки включає в себе згортку зображень, виділення контурів, фільтрацію зображення, побудову гістограм.

Умовно всі алгоритми можна розділити на групи.

2.3.1.1 Фільтрація і знаходження контуру

Всі СОРО використовують ті чи інші види фільтрів і алгоритми знаходження контурів об'єктів. У таких системах широко використовуються стандартні алгоритми, такі як фільтри Собеля, Калмана і Конні, скелетизації, двовимірні вейвлет перетворення, алгоритм мурашки та інші. В окремих системах використовуються і оригінальні алгоритми, кожен з яких змінено відповідно до конкретного завдання. У кожному проекті розробники виходять з наявної апаратної бази і евристичних міркувань, зазвичай без переконливої аргументації свого вибору.

2.3.1.2 Гістограми яскравості і контурів

Одночасно з алгоритмами знаходження контурів застосовуються також гістограми. У роботах описані гістограми двох видів: гістограми яскравості зображення і гістограми контурів. Як правило, застосовують ці алгоритми паралельно.

2.3.1.3 Виділення текстурних параметрів і сегментація

Текстури є однією з найбільш корисних інтегральних характеристик зображень. Тому в деяких СОРО використовується виділення текстур. Текстура зазвичай включає або впорядковані візерунки, що складаються з правильних елементів, або велику кількість дрібних предметів (трава, листя дерев, гравій та ін.). Розбиття зображення на ділянки з однаковою текстурою називається сегментацією текстури. Поряд з такою сегментацією можна використовувати сегментацію кольору, автоматично розбиваючи зображення на області когерентного кольору. У багатьох випадках таке розбиття допомагає вирішити задачу виділення цілісного об'єкта на малюнку. Однак сама по собі сегментація не дозволяє виділити об'єкт, наприклад, якщо об'єкт складається з безлічі кольорових плям (яскравий одяг пішохода).

Найефективніший підхід при вирішенні задачі розбиття зображення на об'єкти – це комбінування текстури та кольору. Серед прикладних систем, що використовують цей підхід, можна виділити систему Blobworld [11]. Для отримання кількох сегментів на зображенні система використовує імовірнісний ОМ-метод (метод очікування-максимізації), який досить успішно справляється із завданням виділення деякого об'єкта на зображенні і відділенні цього об'єкта від фону, але займає досить тривалий час.

2.3.2 Алгоритми розпізнавання

Для розпізнавання об'єктів застосовуються будь-які з доступних класифікаторів. Як зазначалось вище, якість розпізнавання більше залежить від підготовленого коду, ніж від того чи іншого типу класифікатора. Часто в якості класифікатора використовують нейронні мережі. В більшості розглянутих систем це Support Vector Machine (SVM) [12], причому ні в одній роботі немає аргументації, чому слід застосовувати саме цей тип нейронних мереж. Порівняльний аналіз різних класифікаторів також відсутній. Мабуть, такий вибір пов'язаний з тим, що розробники більше уваги приділяють попередній обробці даних, ніж алгоритмам розпізнавання,

просто використовуючи добре розрекламований алгоритм і не вдаючись в дослідження в цій області.

2.4 Огляд існуючих систем

2.4.1 Pre Safe від Mercedes-Benz

Функція виявлення пішоходів розширює функцію гальмування PRE-SAFE®, додаючи виявлення пішоходів та велосипедистів на шляху руху автомобіля.

Система гальмування PRE-SAFE® працює в діапазоні швидкостей автомобіля від 7 до 200 км / год. У випадку нерухомих перешкод, гальмування PRESAFE® активується лише до максимальної швидкості 72 км / год і реагує на:

а) перешкоди, наприклад, зупинки або припарковані транспортні засоби на шляху автомобіля;

б) пішоходів на шляху автомобіля;

Коли PRE-SAFE® Brake реєструє ризик зіткнення з перешкодами в діапазоні виявлення радіолокаційної датчикової системи або стерео-камери, панель приладів забезпечує візуальне або візуально-акустичне попередження. Якщо водій не реагує (ухиляється від маневру), починається автономне часткове гальмування. Якщо від водія досі немає реакції на це, виконується аварійна зупинка, і транспортний засіб зупиняється, якщо всі вимоги виконані. У кращому випадку нещасних випадків можна уникнути зі швидкістю до 50 км / год [13].

Стереофункціональна камера складається з двох моно-камер з високою роздільною здатністю, розміщених приблизно на відстані 20 сантиметрів в одному корпусі. Хоча монокамера лише оцінює відстань до об'єктів, багатофункціональна стереокамера може вимірювати відстань до об'єкта, а також висоту об'єкта над дорогою. Це стало можливим завдяки перспективній різниці між лівим і правим оптичними шляхами. Електроніка оцінки в стереофункціональній камері використовує той самий ефект, що дозволяє людям бачити в трьох вимірах: зміщення паралакса між двома зображеннями. При помірній відстані від 20 до 30 метрів стерео-багатофункціональна камера може розрахувати відстань до об'єкта з точністю від 20

до 30 сантиметрів. Висока роздільна здатність також важлива при складних умовах; наприклад, коли кілька об'єктів дуже близькі один до одного, коли об'єкти частково закриті, або коли контраст між об'єктом і фоном низький. Важливою властивістю стереофункціональної камери є її здатність розрізняти тіні від реальних об'єктів і дорожню розмітку.

Як і тривимірне розташування виявленого об'єкта, стереофункціональна камера також надає важливу додаткову інформацію: для кожного пікселя виявленого об'єкта можна обчислити напрямок руху по горизонтальній, вертикальній і поздовжній осях. Це чітко показує, чи рухається об'єкт. У поєднанні з класифікацією об'єктів (наприклад, пішохода) за допомогою загальних ознак досягається високий ступінь визначеності. Можна, наприклад, розрахувати точну точку зіткнення а, отже, найкраще використовувати залишковий часовий інтервал для виконання захисних заходів. Стерео-багатофункціональна камера допомагає у всьому діапазоні швидкості автомобіля.

Крім того, використовуються інші датчики:

- а) передній дальній радарний датчик;
- б) радар короткого діапазону.

В ідеальних умовах передній радіолокаційний датчик далекого радіуса може виявляти об'єкти на відстані 200 м. На рисунку 2.5 показано розташування датчиків системи. На рисунку 2.6 показано приклад роботи системи виявлення та класифікації об'єктів [13].

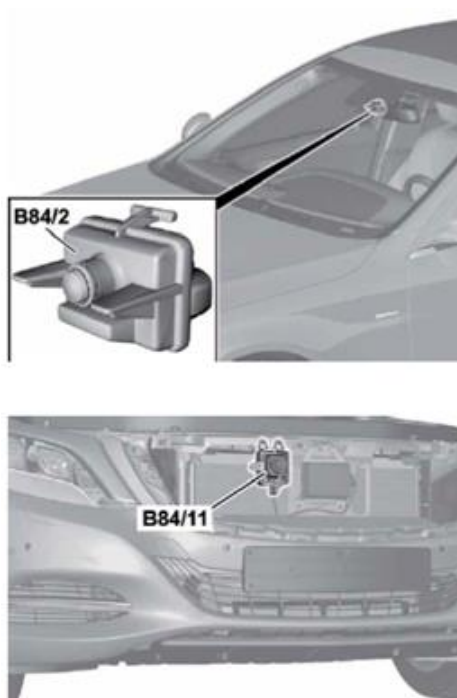


Рисунок 2.5 – Розташування датчиків системи PRE-SAFE [13]



Рисунок 2.6 – Робота системи виявлення [13]

2.4.2 Pre-collision system від Toyota

Система Pre-collision використовує радіолокаційний датчик та датчик камери для виявлення транспортних засобів та пішоходів перед автомобілем.

Коли система визначить, що можливість лобового зіткнення з транспортним засобом або пішоходом висока, система попередить водія та автоматично підвищить

тиск у гальмівній системі для переведення системи у режим максимальної готовності, щоб допомогти водієві уникнути зіткнення.

Попередження переднього зіткнення: швидкість транспортного засобу становить приблизно від 10 до 180 км / год.

Для виявлення пішохода швидкість транспортного засобу становить приблизно від 10 до 80 км / год.

Відносна швидкість між транспортним засобом та транспортним засобом або пішоходом попереду – приблизно 10 км / год або більше. На рисунку 2.7 показано розташування датчиків системи [14].



Рисунок 2.7 – Розташування датчиків системи Pre-Collision [14]

Система Pre-collision виявляє пішоходів на основі розміру, статури та руху виявленого об'єкта. Однак, пішохода може бути не виявлено залежно від навколишнього освітлення, положення тіла та інших факторів.

Toyota із системою Pre-collision поставляється з передньою камерою та невидимим лазером. Система використовує ці компоненти для вимірювання відстані між автомобілем та будь-якою перешкодою. Враховуючи швидкість, система визначає ймовірність зіткнення [14].

У транспортних засобах із виявленням пішоходів додаткова система, встановлена на решітці, використовує радіолокатор міліметрової довжини хвилі для сканування оточення на предмет менших об'єктів, які Pre-collision може пропустити.

2.4.3 Pre-Sense від Audi

Автономна система аварійного гальмування Pre-sense використовує радар та монокулярну камеру та була представлена в 2010 році на Audi A8. Pre-sense plus працює в чотири етапи. Система спочатку показує попередження про можливу аварію, активізуючи попереджувальні сигнали про небезпеку, закриваючи вікна та люк, а також використовує попередній натяг ременів безпеки. Попередження супроводжується легким гальмуванням для привертання уваги водія. Третя фаза ініціює автономне часткове гальмування з прискоренням 3 м/с^2 . Четверта фаза збільшує швидкість гальмування до 5 м/с^2 з подальшим автоматичним повним гальмуванням, приблизно за півсекунди до прогнозованого удару [15].

2.4.4 Driving Assistant від BMW

У 2013 році Driving Assistant Plus був представлений на більшості моделей. Система складається з камери, встановленої за решіткою радіатора, системи попередження про виїзд на зустрічну смугу руху та в деяких випадках з передніми радіолокаційними датчиками для виявлення транспортних засобів попереду. Якщо водій не відреагує на попередження про можливе зіткнення, система поступово збільшить гальмівний тиск і при необхідності застосує повне гальмування. У разі аварії система може привести автомобіль у стан максимальної готовності. Пізніші ітерації системи на автомобілях, оснащених системою автоматичного круїз-контролю, вдосконалюються за рахунок поєднання радіолокаційного та камерного виявлення під час туману, дощу та інших ситуацій, коли нормальна робота камери може бути порушена [16]. На рисунку 2.8 показано приклад роботи системи.



Рисунок 2.8 – Приклад роботи системи Driving Assistant Plus [16]

2.4.5 EyeSight від Subaru

Система Subaru під назвою EyeSight була представлена у 2008 році. Вона використовує стереокамери для виявлення пішоходів та велосипедистів. EyeSight дозволяє гальмувати перед зіткненням та використовувати адаптивний круїз-контроль на всіх швидкостях. Сигналізація використовується для попередження водія про потенційну небезпеку. Контроль гальмування перед зіткненням був удосконалений у 2010 році, щоб дозволити автомобілю автоматично зупинитися, якщо різниця швидкостей між обладнаним EyeSight транспортним засобом та об'єктом попереду становить менше 30 км/год. При швидкості вище 30 км / год автомобіль автоматично знижує свою швидкість. Це також дозволяє автомобілю ввімкнути гальмування, якщо існує ризик лобового зіткнення і водій раптово застосує гальмо. Різниця швидкостей для автоматичної зупинки була підвищена до 50 км/год в 2013 році за допомогою встановлення покращених камер. У 2010 році також було вдосконалено адаптивний круїз-контроль, щоб дозволити автоматичне гальмування в аварійному режимі, повністю зупинивши автомобіль. У 2013 році до камер було додано кольорову камеру, що дозволяє системі розпізнати гальмівні ліхтарі та габаритні ліхтарі. Subaru також додав систему активного ведення смуги руху (утримання транспортного засобу посередині смуги руху та застосування рульової

сили для утримання транспортного засобу в смузі руху при ненавмисному перетині маркерів смуги руху) та управління дроселем (для запобігання раптового ненавмисного прискорення вперед та назад) [17]. Приклад роботи системи показано на рисунку 2.9.

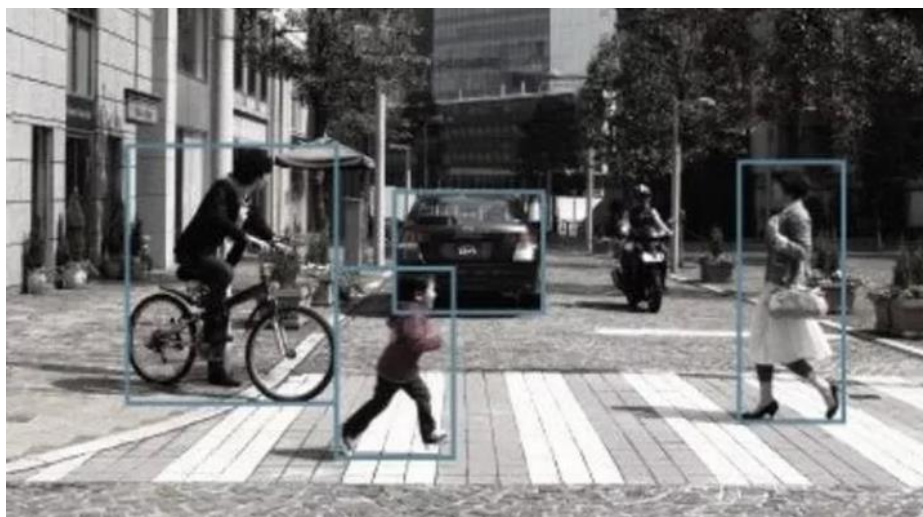


Рисунок 2.9 – Приклад роботи системи EyeSight [17]

2.4.6 Tesla Autopilot

Tesla Autopilot – це вдосконалена система допомоги водієві, запропонована Tesla, яка має можливість утримання у смузі руху, адаптивний круїз-контроль, можливість самостійного паркування, можливість автоматичної зміни смуг руху та можливість викликати автомобіль до гаража чи місця для паркування.

Основна мета – створення повністю автономного авто (FSD) [18].

Перша версія, що доступна для клієнтів, була випущена в 2015 році як частина версії Tesla 7.0. Версія програмного забезпечення 7.1 потім вилучила деякі функції, щоб відмовити клієнтів від ризикованої поведінки та додала технологію віддаленої парковки Summon, яка може переміщувати машину вперед та назад під віддаленим людським управлінням без водія в автомобілі [19].

У 2016 році Tesla заявила, що всі транспортні засоби поставляються з необхідним зондуванням та обчислювальним обладнанням, відомим як апаратне забезпечення версії 2 (HW2) для майбутнього повного самостійного водіння.

У поліпшеному автопілоті є такі часткові здатності до автономного водіння: автоматична зміна смуги руху, перехід з однієї смуги в іншу, вихід зі смуги.

У 2019 році Tesla почала випускати оновлення для навігації на автопілоті, яке не вимагає підтвердження зміни смуги руху, але вимагає від водія тримати руки на кермі. Автомобіль буде самостійно орієнтуватися на дорозі, але водієві потрібно контролювати рух. Ця можливість доступна для тих, хто придбав розширений автопілот.

У 2019 році Tesla запропонувала оновлений автопілот в Європі, щоб прийняти нове положення ООН / ЄЕК R79, пов'язане з функцією автономного керування.

Потім Tesla випустила автопілот V10. Для автопілоту було розроблено великі нейронні мережі, які не можуть бути використані через відсутність обчислювальних ресурсів у поточному обладнанні Tesla. Наступна версія апаратного забезпечення (3.0) надає ресурси для забезпечення більшої точності прогнозів.

HW 3.0 включає в себе спеціальний процесор, розроблений Tesla, використовуючи чіпи для інтегральних мікросхем (ASIC). Tesla заявила, що нова система буде обробляти 2300 кадрів в секунду, що є 21-кратним покращенням обробки зображень порівняно з апаратним забезпеченням 2.5, здатним обробляти до 110 кадрів в секунду. Фірма описала його як "прискорювач нейронної мережі".

Tesla стверджує, що HW 3.0 має покращену продуктивність у порівнянні з HW 2.5 у 2,5 рази та зменшення вартості у 0.8 рази; обладнання має 12-ти ядерний процесори ARM Cortex-A72 з частотою 2,2 ГГц, 2х NNA (прискорювачі нейронної мережі) з частотою 2 ГГц і графічний процесор у Mail GPU 1 ГГц [20].

Автопілот може виявити можливе переднє або бічне зіткнення з іншим транспортним засобом, велосипедом або пішоходом на відстані 160 м. Автопілот має автоматичне екстрене гальмування, яке виявляє потенційно небезпечні об'єкти. Автомобіль також може автоматично змінити траєкторію, щоб уникнути зіткнення. Автопілот також може автоматично регулювати габаритні вогні, світло ближніх та

дальніх фар в залежності від зміни освітлення в нічний час доби. Приклад роботи системи показано на рисунку 2.10.



Рисунок 2.10 – Приклад роботи системи Tesla Autopilot [20]

2.5 Дослідження IIHS

Зменшення аварій пов'язаних з пішоходами є метою нових досліджень IIHS у галузі автоматичних систем екстреного гальмування, які можуть виявляти пішоходів та тварин. У перших випробуваннях автомобілів 2018-19 років із цією функцією 9 з 11 оцінених малих позашляховиків отримали високі та добрі оцінки.

2018–19 Honda CR-V, 2019 Subaru Forester, 2019 Toyota RAV4 та Volvo XC40 2019 року показали найкращі результати і отримали оцінку відмінно. П'ять моделей отримали оцінку добре. Це Chevrolet Equinox 2019, 2018–19 Hyundai Kona, 2019 Kia Sportage, 2018–19 Mazda CX-5 та 2019 Nissan Rogue. Mitsubishi Outlander в 2019 році отримав оцінку задовільно, а BMW X1 2018–19 років отримала оцінку погано [21].

Аналіз HLDI 2018 року встановив, що система EyeSight Subaru із виявленням пішоходів знизила ставку ймовірних вимог щодо страхування пішоходів на 35 відсотків порівняно з тими ж транспортними засобами без системи.

Взагалі, системи виявлення пішоходів використовують монокамери, спрямовані вперед або стереокамери, встановлені біля дзеркала заднього огляду, а

також радіолокаційні датчики на передній решітці автомобіля для постійного сканування проїжджої частини та горизонту і, в деяких випадках, велосипедистів чи тварин, які можуть перетнути дорогу. Алгоритми класифікують об'єкти як людей, велосипедистів чи тварин, прогнозують їх переміщення та визначають швидкість руху транспортного засобу стосовно них. Якщо зіткнення неминуче, система, як правило, попереджає водія і може застосувати гальмування набагато швидше, ніж людина могла б зреагувати.

Тести стосуються трьох загальних сценаріїв аварії з пішоходами.

Перший сценарій стосується дорослого пішохода, що перетинає дорогу з правого боку. Це найпоширеніший тип аварії з участю пішохода. Схематичне зображення сценарію показано рис. 2.11.

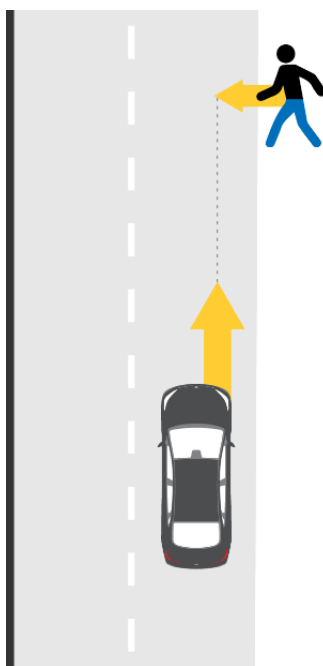


Рисунок 2.11 – Схематичне зображення першого сценарію [21]

Другий тест імітує дитину, що виходить на дорогу із-за двох припаркованих транспортних засобів. Схематичне зображення сценарію показано рисунку 2.12.

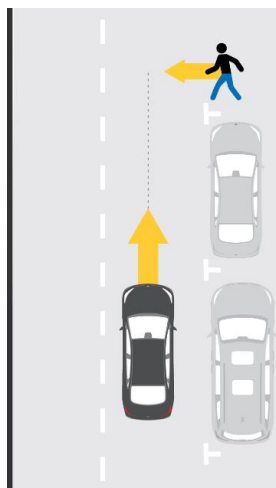


Рисунок 2.12 – Схематичне зображення другого сценарію [21]

Третій тестовий сценарій повторює дорослу людину, яка рухається вздовж дороги. Схематичне зображення сценарію показано на рисунку 2.13.

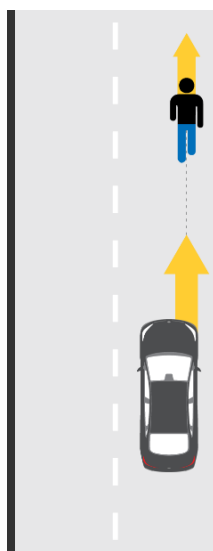


Рисунок 2.13 – Схематичне зображення третього сценарію [21]

Транспортні засоби оцінюються відповідно до їх середнього зниження швидкості у п'яти спробах на сухому покритті. Випробування проводяться зі швидкістю 20 км/год і 50 км/год за першим та другим сценаріями, а також при 50 км/год і 60 км/год при третьому сценарії.

Чотири авто, що отримали оцінку відмінно та п'ять позашляховиків, що отримали оцінку добре показали добрі результати в кожному сценарії. Це означало,

що позашляховики майже уникали, а в деяких випадках уникали ударів пішохідних муляжів.

Найскладніший тест – сценарій, де дитина виходить із-за припаркованих авто. Манекен висотою 115 см, який представляє 7-річну дитину середнього розміру, схований автомобілем та позашляховиком, припаркованих з правого боку дороги. Для камери немає чіткої лінії огляду, поки манекен не з'явиться із-за припаркованих транспортних засобів. Коли манекен виходить на смугу проїзду, випробувальний автомобіль має приблизно 1,5 секунди до зіткнення.

Лише Forester та RAV4 уникали ударів по манекенам у кожному тесті за 1 та 2 сценаріями. BMW X1, який поставляється із системою виявлення пішоходів, зовсім не гальмував у 3 сценарії. Цей позашляховик мав мінімальне зниження швидкості в інших тестах.

Аналіз IIHS 2011 року за даними аварій 2005 – 2009 років виявив, що системи виявлення пішоходів можуть потенційно пом'якшити або запобігти до 65 відсотків аварій транспортних засобів з пішоходами у трьох найпоширеніших конфігураціях аварій та зменшити на 58 відсотків кількість загиблих пішоходів у цих аваріях.

Приблизно дві третини систем попередження аварій, що пропонуються на моделях 2019 року, мають можливості виявлення пішоходів. Багато з них також можуть виявляти та реагувати на велосипедистів, а в деяких випадках і на тварин, хоча IIHS не оцінював ці можливості.

2.6 Висновки

У розділі приведено загальну характеристику систем оптичного розпізнавання, їх класифікації за різними показниками. Також, приведено загальну характеристику та опис алгоритмів попередньої обробки та розпізнавання. Представлено огляд існуючих систем розпізнавання та систем автономного управління. В кінці розділу наводяться дані про дослідження систем екстреного гальмування страхового інституту безпеки дорожнього руху за 2019 рік.

3 ОПИС ТА МАТЕМАТИЧНА МОДЕЛЬ АЛГОРИТМІВ

3.1 Загальна характеристика

В останні роки, автономні автомобілі стали основним напрямком роботи для підвищення безпеки на дорогах загального користування. В свою чергу, головним завданням при розробці таких транспортних засобів залишається швидке та якісне розпізнавання об'єктів.

Точність класифікації, виявлення та оцінки позицій ґрунтуються на якості інформації, що надходить з різних даччиків інформації. Одним з головних завдань, що вирішується у роботі є надання огляду сучасних методів виявлення об'єктів та порівняння існуючих алгоритмів.

Виявлення об'єктів за допомогою кольорової камери включає в себе як класифікацію, так і локалізацію кожного об'єкта, який представляє інтерес. Заздалегідь невідомо скільки об'єктів очікується знайти в кожному зображенні, що означає, що для кожного вхідного зображення існує різна кількість виходів. Також, невідомо, де ці об'єкти можуть розташовуватися на зображенні або якими можуть бути їх розміри. В результаті, постає досить глибока проблема – розпізнавання об'єктів.

З розвитком глибинного машинного навчання (Deep Learning) методи виявлення об'єктів за допомогою глибинного навчання значно перевершили багато традиційних методів як в точності, так і в швидкості. На основі методів глибинного навчання, існують системи, що покращують результати виявлення традиційними методами.

Загалом, є два підходи до виявлення об'єктів на зображеннях за допомогою глибинного навчання. Один із підходів базується на основі аналізу регіонів, наприклад, Faster-RCNN. Цей метод спочатку обробляє все вхідне зображення за допомогою кількох згорткових шарів, щоб отримати карту зображень. Тоді, окрема мережа використовує ці згорткові функції, щоб запропонувати можливі регіони для виявлення. Нарешті, решта мережі надає класифікацію цим запропонованим регіонам. Оскільки в складі є дві частини, то одна використовується для

прогнозування обмежувального прямокутника, а друга саме для класифікації, системи такого типу архітектури можуть значно уповільнити швидкість обробки. Інший підхід використовує одну мережу як для передбачення потенційних регіонів, так і для класифікації міток. Один із прикладів – You only look once (YOLO). YOLO спочатку розділяє зображення на грубі сітки. Для кожної сітки є набір базових обмежувальних боксів. Для кожного базового обмежувального бокса YOLO прогнозує зміщення рівня істинного місця розташування, дає оцінку достовірності та визначає клас, якщо мережа вважає, що в цій сітці є об'єкт. YOLO швидкий, але іноді може давати помилки при виявленні невеликих об'єктів на зображенні.

Алгоритми розпізнавання об'єктів в основному мають однакову послідовність дій. Загальна структура системи виявлення: сенсорна система фіксує дані у вигляді зображення, потім застосовуються методи пропозицій щодо регіонів; отримують функції ознак із запропонованих регіонів і нарешті функції передаються в класифікатор.

3.1.1 Регіони інтересу

Регіони інтересу (Region of Interest або ROI), також відомі як регіональна пропозиція, вважаються першими і найважливішими кроками у системах виявлення. Для зручності пошуку на цьому етапі застосовуються деякі методи обробки зображень.

Такі характеристики, як краї, лінії, візерунки, витягуються та обробляються з використанням класифікаторів для визначення класу об'єкта. Підходи, які можуть бути використовуватися для пошуку регіонів інтересу – послідовно переміщуване вікно, селективний пошук та локально декоррельований канал характеристик (locally decorrelated channel features або LDCF).

При використанні метода послідовно переміщуваного вікна, зображення сканується за допомогою невеликих квадратних регіонів, що можуть мати різні розміри, а також сканування може проводитися в будь-якому напрямі – горизонтальному чи вертикальному [22].

Жодні позитивні регіони не відкидаються, оскільки всі регіони подаються до нейронної мережі для подальшої обробки. Це забезпечує високу точність але з більш високим рівнем обчислювальної складності через велику кількість помилково позитивних результатів.

Вибірковий пошук використовує грубий фільтр для виявлення незалежних від класу регіонів. Такий підхід успішно використовується зі згортковими нейронними мережами для вилучення та класифікації ознак. Підхід зменшує кількість запропонованих регіонів, зменшуючи обчислювальні витрати.

LDCF може виявляти різні об'єкти з високою точністю. Для подальшого вдосконалення цей підхід поєднується з нейронною мережею, де обробляється велика кількість регіонів, кожен з яких має певне значення довіри. Значення довіри відноситься до ймовірності того, що в кадрі міститься певний об'єкт. Це забезпечує компроміс точності та коефіцієнту корисної дії детектора [24].

Підхід з використанням послідовно переміщуваного вікна – найпростіша техніка, і вона пристосована для використання зображень з різними співвідношення сторін та масштабами. Однак більш складні алгоритми можуть призвести до зниження кількості знайдених регіонів, збільшуючи кількість помилково-позитивних результатів.

3.1.2 Отримання ознак

Алгоритми для отримання ознак та класифікації для знаходження об'єктів можуть бути створені власноруч або можуть бути використані методи на основі алгоритмів глибинного машинного навчання. Методи, що написані власноруч засновані на моделях, що розроблені для низькорівневих функцій. Ці моделі були розроблені на таких методах як гістограми орієнтованих градієнтів (HOG).

Гістограма орієнтованих градієнтів – це дескриптор функції, що використовується для систем комп'ютерного зору та обробки зображення для розпізнавання об'єктів. Техніка алгоритму полягає в розрахунку входжень градієнтної орієнтації в локалізованих частинах зображення. Такий метод

аналогічний до методу гістограми орієнтації ребер, дескрипторам перетворення об'єктів, але відрізняється тим, що він розраховується на щільній сітці із рівномірно розташованих комірок і використовує перехресну нормалізацію локального контрасту з метою підвищення точності [24].

Головна ідея такого алгоритму полягає в тому, що зовнішній вигляд та форму локального об'єкту на зображенні можливо описати за допомогою розподілення градієнтів інтенсивності або направленості меж об'єкта. Зображення поділяється на невеликі зв'язані області, що називаються комірками, а потім для пікселів в кожній комірці складається гістограма напрямлень градієнта. Дескриптор – це об'єднання отриманих гістограм. Для підвищення точності локальні гістограми можна нормалізувати за контрастом, для цього потрібно розрахувати міру інтенсивності в більшій частині зображення, яку називають блоком, а потім використати отримане значення для нормалізації усіх комірок у блоці. Така нормалізація призводить до кращої інваріантності відносно змін освітленості та затемнення.

Дескриптор гістограм орієнтованих градієнтів має декілька вагомих переваг в порівнянні з іншими видами дескрипторів. Оскільки він застосовується до локальних комірок, то він має більшу інваріантність до геометричних та фотометричних перетворень, крім орієнтації об'єкту. Такі зміни будуть з'являтися тільки в великих просторових областях. Крім того, груба просторова вибірка, точна орієнтаційна вибірка та сильна локальна фотометрична нормалізація дозволяє виключати рух окремих об'єктів, таких як пішоходи, поки вони підтримують вертикальне положення. Таким чином, можна зробити висновок, що дескриптор гістограм орієнтованих градієнтів найбільш вдало підходить для систем розпізнавання та трекінгу пішоходів.

Першим етапом розрахунку в багатьох детекторах ознак при попередній обробці зображень є забезпечення нормалізації значень кольорів та гами. Однак, цей крок може бути пропущений при розрахунку дескрипторів гістограм орієнтованих градієнтів, оскільки в результаті наступної нормалізації дескриптора буде досягнуто того ж самого результату. А отже, попередня обробка зображення має невеликий вплив на якість та час розрахунку. Замість цього, першим кроком роботи алгоритму є

обчислення значень градієнтів. Найбільшою популярністю користується метод 1-D центрованої точкової дискретної похідної маски в одному або обох із вертикального або горизонтального напрямків. Однак, цей метод потребує фільтрації даних кольору або інтенсивності зображення з наступними ядрами фільтрів:

$$[-1, 0, 1] \text{ та } [-1, 0, 1]^T. \quad (3.1)$$

Автори існуючих робіт протестували також і інші, більш складні маски, такі як маска Собеля 3x3, або діагональні маски, але ці маски при розпізнаванні пішоходів, як правило, працюють гірше. Також було проведено дослідження з використанням Гаусівського згладжування перед використанням похідної маски та встановлено, що пропуск будь-якого згладжування краще виконується на практиці.

Другий крок розрахунку – це створення гістограм комірок. Кожен піксель в комірці виконує зважуване голосування для каналу гістограми на основі орієнтації, що базується на значеннях, знайдених при обчисленні градієнту. Самі комірки можуть бути будь-якої форми – прямокутної або радіальної, а канали гістограми рівномірно розподілені від 0° до 180° або від 0° до 360°, в залежності від знаку градієнту. Було помічено, що непідписані градієнти, що використовуються з дев'ятьма каналами гістограми, показали найкращі результати в експериментах з розпізнаванням людей. Що стосується ваги голосування, то внесок в пікселі може бути або величиною самого градієнта, або деякою функцією величини. В тестах сама величина градієнта зазвичай дає найкращі результати. Інші варіанти ваги голосування можуть включати квадратний корінь або квадрат величини градієнта або деяку урізану частину величини [24].

Аби врахувати зміну освітленості та контрасту, сили градієнту повинні бути локально нормалізованими, що потребує групування комірок разом в більш великі, просторово зв'язані блоки. Тоді дескриптор гістограм орієнтованих градієнтів є зчепленим вектором компонентів нормалізованих гістограм комірок із усіх областей блоку. Ці блоки зазвичай перекриваються, що означає, що кожна комірка вносить внесок в кінцевий дескриптор більше одного разу. Існують дві основні геометрії

блоків: прямокутні блоки R-HOG і круглі блоки C-HOG. Блоки R-HOG, як правило, представляють собою квадратні сітки, побудовані трьома параметрами: кількістю осередків на блок, кількістю пікселів на осередок і кількістю каналів на гістограму осередку. В експерименті з виявлення людини були знайдені оптимальні параметри: чотири осередки 8x8 пікселів на блок (16x16 пікселів на блок) з дев'ятьма каналами гістограми. Більш того, вони виявили, що деяке незначне поліпшення продуктивності може бути досягнуто шляхом застосування Гаусівського просторового вікна в кожному блоці перед табулюванням голосів гістограми, щоб менше зважувати пікселі навколо краю блоку. Блоки R-HOG виглядають дуже схоже на дескриптори масштабно-інваріантного перетворення ознак (SIFT); однак, незважаючи на їх однакове формування, блоки R-HOG обчислюються в щільних сітках в деякому єдиному масштабі без вирівнювання орієнтації, тоді як дескриптори SIFT зазвичай обчислюються в розріджених, незмінних за масштабом точках зображення ключа і повертаються для вирівнювання орієнтації. Крім того, блоки R-HOG використовуються спільно для кодування інформації про просторові форми, тоді як дескриптори SIFT використовуються окремо.

Круглі блоки HOG (C-HOG) можуть бути знайдені в двох варіантах: з одним центральною осередком і з центральним кутовий осередком з кутовим розділенням. Крім того, ці блоки C-HOG можуть бути описані з чотирма параметрами: кількість кутових і радіальних блоків, радіус центрального блока і коефіцієнт розширення для радіуса додаткових радіальних блоків. Також було виявлено, що два основні варіанти забезпечують однакову продуктивність, і що дві радіальні комірки з чотирма кутовими осередками, центральним радіусом 4 пікселя і коефіцієнтом розширення 2 забезпечують кращу продуктивність в експериментах. Крім того, зважування по Гаусу не дає користі при використанні в поєднанні з блоками C-HOG. Блоки C-HOG виглядають аналогічно дескрипторам контексту форми, але сильно відрізняються тим, що блоки C-HOG містять осередки з декількома каналами орієнтації, в той час як контексти форми використовують тільки один канал присутності ребер у своєму формулюванні.

Методи глибинного машинного навчання дозволяють мережі визначати функції, які неможливо реалізувати власноруч. Це може забезпечити більш високий рівень абстракції. Залежно від застосування можуть використовуватися різні методи. Для кожної вхідної області створюється вектор дійсних або довічних значень. Вихідний вектор являє собою видимі характеристики пропонованих регіонів.

3.1.3 Класифікація

Вихідний вектор, отриманий на етапі виявлення функції, подається в класифікатор, щоб визначити, чи існує об'єкт у запропонованих регіонах у вигляді бінарної мітки. Класифікатори, про які йшла мова до цього використовували екстрактори функцій, включаючи AdaBoost та SVM. Однак із розвитком машинного навчання частіше реалізуються підходи на основі згорткових нейронних мережах (CNN) для класифікації. Ці підходи на основі згорткових нейронних мереж будуть описані далі.

Підмножина штучного інтелекту та машинного навчання – глибоке машинне навчання (DL) було вперше представлено у 90-х роках, але лише нещодавно ця галузь набула широко вжитку завдяки великій підтримці та зменшенню витрат на обчислювальну техніку (наприклад, одиниці графічної обробки (GPU)) та ефективні алгоритми навчання. Зокрема, алгоритми згорткових нейронних мереж (CNN) знайшли використання у галузі комп'ютерного зору та аналізу зображень для виявлення об'єктів, класифікації зображень та розпізнавання облич.

Згорткові нейронні мережі (CNN) – це тип технології глибинного машинного навчання, який має високу ефективність у багатьох сферах як у якості алгоритму розпізнавання, так і класифікації об'єктів. Надійність мереж CNN впливає з того, що вони здатні отримувати інформацію з вмісту неопрацьованих пікселів і автоматично вивчати функції. Така мережа реалізує такі можливості виконуючи різні операції, як правило, комбінація фільтрації, об'єднання та нелінійна активація. Одна з переваг використання CNN для вилучення характеристик, порівняно з іншими методами,

полягає в тому, що CNN отримують характеристики із зображень без явного програмування [25].

З 2012 року для розпізнавання об'єктів були розроблені нові підходи, засновані на методах глибинного навчання, наприклад, AlexNet, мережа CNN розроблена Алексом Крижевським і названа на честь розробника [26].

AlexNet пройшов навчання з використанням набору даних ImageNet. Для ImageNet два типи показників ефективності: top-1 (з використанням повного набору тестів) і top-5 (з використанням неповного тестового набору). Показники помилок AlexNet склали 37,5% і 17,0% для top-1 та top-5 відповідно. До результатів AlexNet найкраща ефективність за показниками помилок становила 47,1% та 28,2%. Ці результати допомогли розробити обладнання для підвищення ефективності роботи CNN з метою підвищення точності виявлення [26].

Нейронні мережі глибинного машинного навчання використовують кілька шарів, які здатні отримувати характеристики, такі як краї або візерунки на зображеннях, і використовують ці характеристики для класифікації об'єкта. Таким чином, глибинні нейронні мережі, такі як CNN, використовуються для вивчення характеристик розпізнавання об'єктів, таких як пішоходи, автомобілі, тварини. Нейронні мережі із прямим зв'язком складаються з серії обчислювальних вузлів, відомих як нейрони, пов'язані між собою для обробки інформації. Така мережа має назву багат шаровий перцептрон (MLP). Вузли утворюють шари, з'єднані між собою через значення параметрів, які називаються вагами. Нейрон функціонує як класифікатор логістичної регресії. Нейрони використовують нелінійні операції для перетворення вхідних даних та створення межі рішення, в якій дані можуть бути лінійно відокремлювані. Ілюстрацію одиничного перцептрона показано на рисунку 3.1. Кілька шарів цих перцептронів створюють MLP або нейронну мережу (рисунок 3.2). Нейронна мережа на рисунку 3.2 є повністю зв'язаною мережею. Це означає, що кожен нейрон отримує вхід від кожного нейрона з попереднього шару.

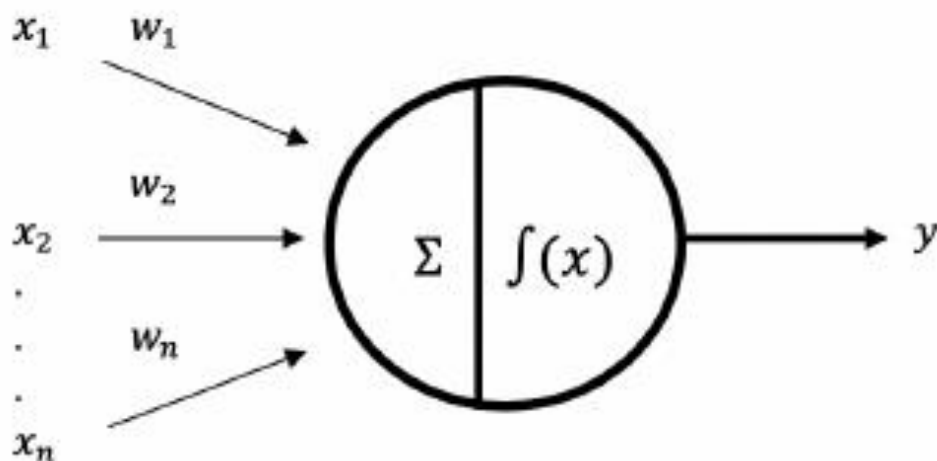


Рисунок 3.1 – Одиничний перцептрон [27]

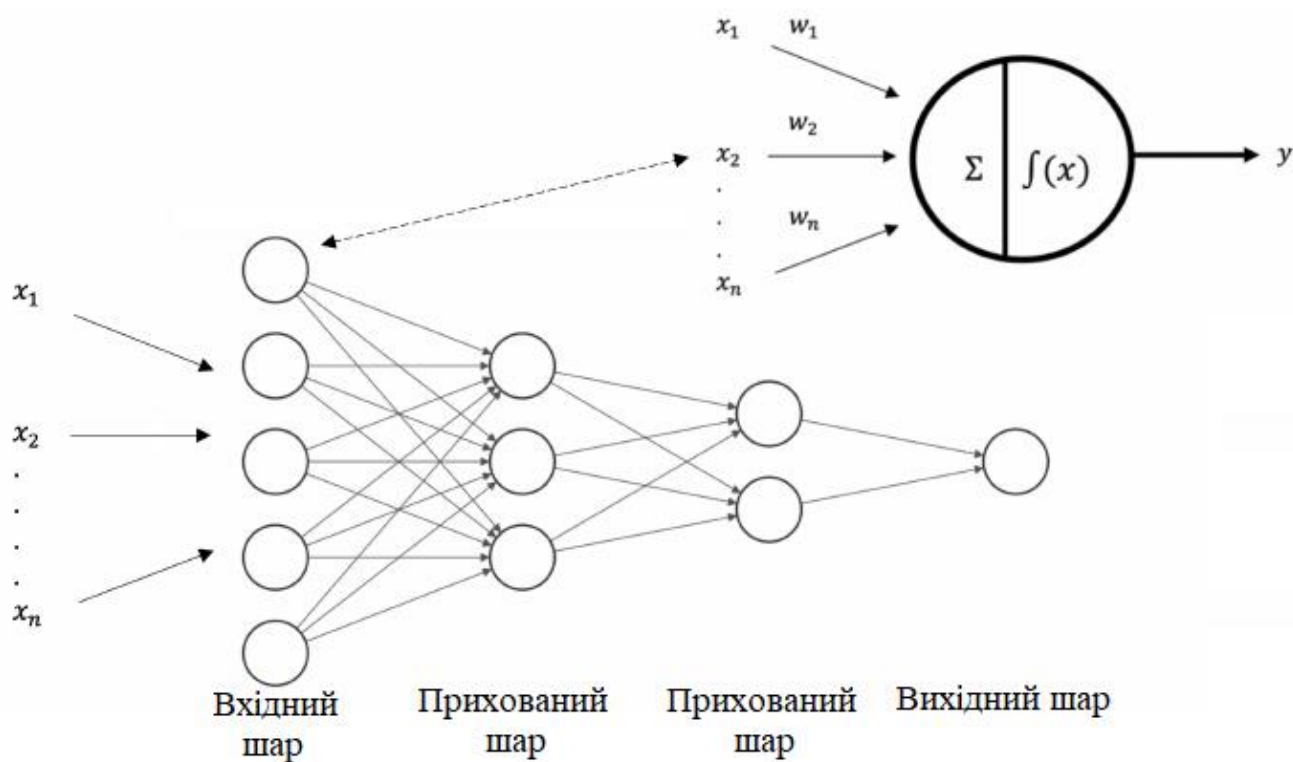


Рисунок 3.2 – Повністю зв'язана нейрона мережа [27]

Для CNN існують шари згортки всередині прихованих шарів для проведення обчислень згортки.

Метою глибокого машинного навчання є виявлення об'єктів в одному або декількох кадрах, подібно до того, як люди виявляють і взаємодіють з об'єктами. З

сучасним розвитком програмного та апаратного забезпечення відбувся справжній прогрес у галузі комп'ютерного зору. Існує багато матеріалів щодо систем розпізнавання та систем стеження, включаючи сенсорні технології та методи обробки. Однак, з появою нових алгоритмів машинного навчання потрібно проводити сучасний аналіз алгоритмів.

3.2 Згорткові нейронні мережі

До впровадження сучасних нейронних мереж основні нейронні мережі (такі як на рисунку 3.2) не мали змоги отримати корисні функції із необроблених даних із сенсорів.

Для пошуку суттєвих особливостей використовувались методи по типу гістограми орієнтованих градієнтів.

Щоб подолати це та підвищити продуктивність нейронної мережі, були розроблені згорткові нейронні мережі (рисунок 3.3). Згорткові мережі засновані мережах з прямим зв'язком (де вихід нейрона був би входом іншого набору нейронів у наступному шарі). CNN використовують згорткові операції для отримання характеристик об'єктів із вхідних даних (наприклад, зображень, відео); при цьому використовуючи різні ядра на кожному шарі. Значення активації нейрону в шарах представляють відфільтровані вхідні дані. Обробляються різні регіони вхідних даних, використовуючи при цьому згорткові операції для виявлення шаблонів у даних. Після генерування карти характеристик виконуються згорткові операції для всіх вхідних даних. Карта характеристик – це подання активації різних частин зображення. Вона використовується для встановлення параметризації ваги та зміщення шарів, що дозволяють вивчити особливості. Зазвичай максимальний пул використовується після згортання для зменшення розміру входу. Це зменшує вимоги до обчислення так як параметри входу зменшуються [28].

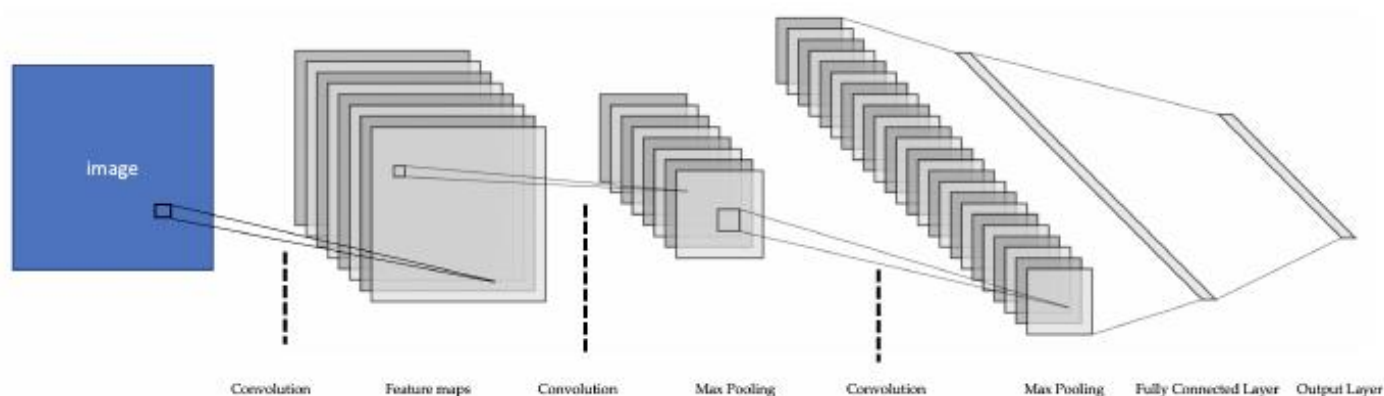


Рисунок 3.3 – Архітектура згорткової мережі [28]

Використовуваний згортковий оператор залежить від типу вхідних даних. 2D ядра (тобто фільтри) використовуються для 2D часових послідовностей (наприклад, відео), а 1D ядра використовуються для 1D часових послідовностей. Коли ядро CNN використовується таким чином, вони можуть використовуватися як класифікатори. Маючи декілька шарів, CNN можуть представляти дані в ієрархічному порядку. По мірі того, як шари стають все глибшими, вхідні дані представлені більш абстрактно, щось, що може бути зроблено вручну, наприклад, отримання характеристик, виявиться дуже важким або неможливим. Це дозволило CNN стати більш стандартною практикою у багатьох сферах, таких як комп'ютерне бачення (наприклад, виявлення об'єктів) та розпізнавання тексту. Мережа може автоматично навчитися отримувати корисну інформацію (тобто характеристики) із зображень та кадрів. Оскільки CNN – це техніка глибокого навчання, тут будуть численні нейрони та шари. Кожен шар представляє різні рівні абстракції. Перші кілька шарів вивчають особливості нижчого рівня, такі як краї, криві або візерунки. Глибокі шари поєднують характеристики для ідентифікації об'єктів у рамках. Класифікаційний шар зазвичай складається з одного ряду нейронів. Кількість нейронів залежить від кількості бажаних виходів (тобто, кількості класів). Наприклад, класи можуть бути пішоходи, велосипедисти або автомобілі, а значить, потрібні три класи. Чим вище вихідне значення для одного з цих нейронів класифікатора, тим більший шанс успішного виявлення пішохода або велосипедиста.

Важливо розуміти, що це дає глибокій мережі можливість вивчати характеристики без явного програмування. Вивчена інформація зберігається в межах регульованих параметрів мережі, званих вагами. Для підготовки мережі до вивчення характеристик використовується набір даних. Набір даних складається з великої кількості зображень, що включають об'єкти, який потрібно буде виявити. Таким чином, характеристики витягуються та вивчаються мережею. Однак, оскільки мережа навчається на основі лише наданих наборів даних, її можна обмежити. Тому для розробки більш надійної та точної CNN потрібен дуже великий набір анотованих наборів даних.

Алгоритми глибинного машинного навчання для розпізнавання об'єктів можуть бути однією з наступних двох категорій (рисунок 3.4): двоетапний детектор (підхід щодо пропозиції регіону) або одноетапний (нерегіональний підхід пропозиції). Одноетапний детектор має на меті усунути потребу у вилученні традиційних пропозицій регіону шляхом обробки цих кроків у межах однієї мережі. Одноступеневий детектор може бути простішим у навчанні, з більш високою обчислювальною ефективністю. При такому підході спочатку завершується пропозиція регіонів, а потім глибока мережа проводить класифікацію.

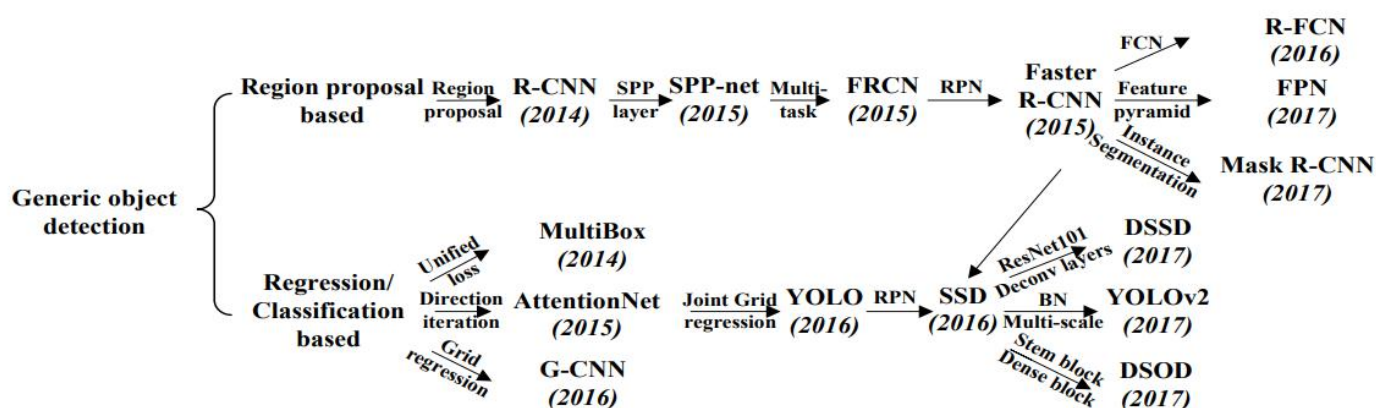


Рисунок 3.4 – Класифікація методів [29]

Методи глибинного навчання, які використовуються для розпізнавання об'єктів, можуть включати пропозицію регіону як частину системи. Деякі з методів,

що базуються на пропозиціях регіону, включають R-CNN, Fast R-CNN та Faster R-CNN [30].

Нерегіональні методи, що базуються на пропозиціях, включають детектор SSD та YOLO [30].

Існують також методи DL щодо нерегіональних пропозицій. Рисунок 3.5 - це детектор, заснований на традиційному регіональному підході і має техніку отримання характеристик, де глибинною мережею є лише класифікатор. На рисунку 3.6 представлена глибока мережа, яка здатна завершити пропозицію регіону та отримати характеристики, а також класифікувати в один крок. Така мережа називається одноетапною. У 2009 році було представлено набір даних Caltech [31] для порівняльного аналізу різних методів виявлення об'єктів. ConvNet (підхід на основі CNN) був представлений у 2013 році з конкурентними результатами порівняно з попередніми методами виявлення об'єктів.

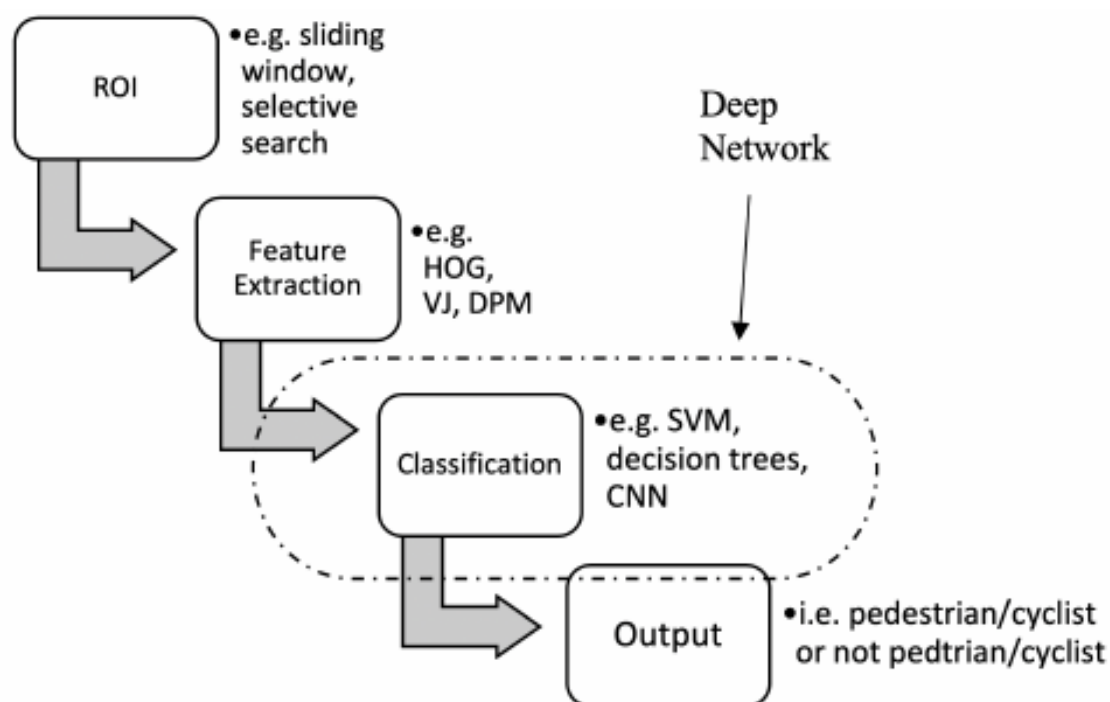


Рисунок 3.5 – Метод, що використовує регіональний підхід [32]

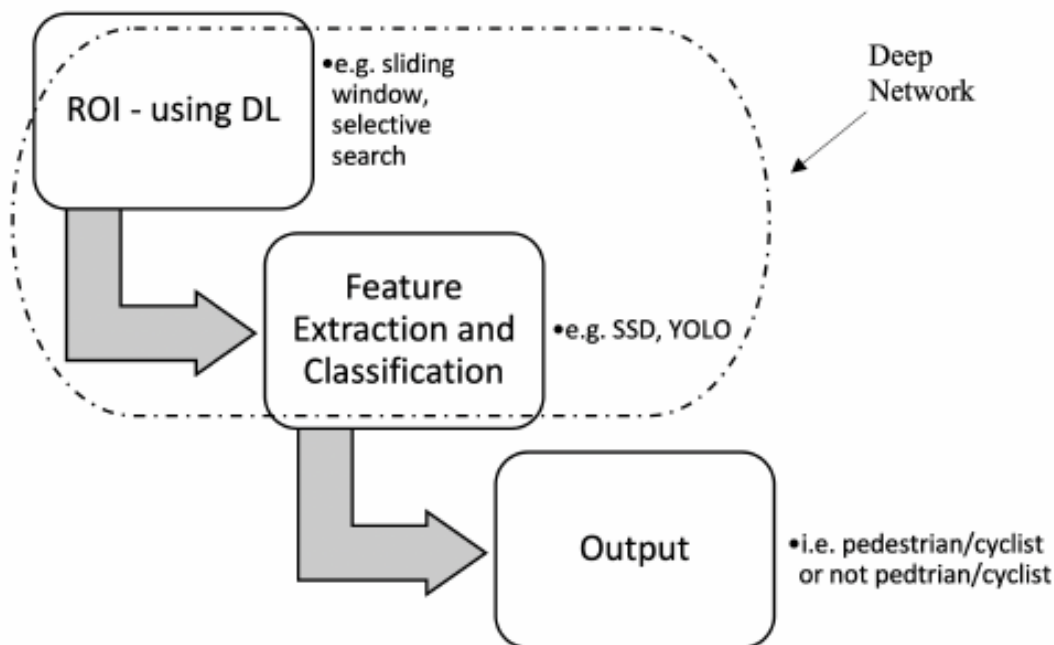


Рисунок 3.6 – Метод, що використовує не регіональний підхід [32]

3.3 Двоетапні детектори

CNN на основі пропозицій регіону (тобто R-CNN, Fast R-CNN) дав позитивні результати для загального підходу виявлення об'єктів. Прикладом такої системи є використання селективного пошуку для генерування регіонів інтересу. Точність цього типу мережі залежить від методики пропозиції регіону, яка застосовується, оскільки ці регіони використовуються для класифікації. Було зроблено підходи до покращення швидкості двоступеневих детекторів, де карта ознак створюються, коли глибока мережа отримує характеристики безпосередньо з регіонів інтересу. Ці методи з тих пір були прийняті, і їх варіанти застосовувались із обнадійливими результатами [33].

Наприклад, у [34] середня точність (AP) досягла вищої оцінки, ніж оцінка KITTI (Технологічний інститут Карлсруе та Технічний інститут Toyota в Чикаго). Отриманий AP збільшився на 9% до 16,7%. З розвитком Fast R-CNN та Faster R-CNN обчислювальна швидкість зросла. R-CNN використовує вибіркового пошук (методика пропозицій щодо регіону) для створення 2000 пропозицій регіону, а не деякої великої

кількості пропозицій регіону. Ці пропозиції подаються в CNN для отримання характеристик, а потім в SVM для класифікації. З цим підходом було кілька питань. А саме, класифікація 2000 пропозицій регіонів потребує значної кількості часу.

Регіональні мережі CNN використовують згортання для зменшення обчислювальних витрат. Однак пропозиція регіону Fast R-CNN не дає змоги використовувати даний тип детектора у режимі реального часу. Для подолання цієї проблеми було запропоновано регіональну мережу пропозицій (RPN), яка ділиться згортковими ознаками з мережею виявлення. Це дозволяє створювати пропозиції регіонів, які майже не мають обчислювальних витрат. RPN - це CNN, яка функціонує, передбачує межі об'єктів (пропозиції регіонів) та одночасно оцінює їх. Це забезпечує детектору якісні пропозиції щодо регіону. Такий підхід дозволяє приблизити показники кадрів в секунду до прийнятної межі для роботи у режимі реального часу, покращуючи якість та точність виявлення об'єктів для загального виявлення об'єктів на основі глибинного навчання [34].

3.3.1 R-CNN

Архітектура мережі R-CNN (Regions With CNN) була розроблена командою з UC Berkley для використання CNN для завдання виявлення об'єктів. Існуючі на цей час підходи для вирішення таких завдань наблизилися до максимальних своїх можливостей, більше вдосконалювати їх показники не було можливості [35].

CNN добре демонструє себе в класифікації зображень, і в даній мережі вони саме через це і використовувалися. Для цього на вхід CNN подавали не все зображення в цілому, а попередньо виділивши іншим способом необхідні регіони, на яких попередньо вважається, що є якісь об'єкти. На даний момент таких підходів було декілька, було обрано селективний пошук, але не наводиться жодних аргументів у використанні саме цього підходу.

У якості CNN використовується готова архітектура – CaffeNet (AlexNet). Такі нейромережі, як і інші, що використовуються для набору зображених ImageNet [36], проводять класифікацію на 1000 класів. R-CNN розроблювалася для виявлення

об'єктів меншої кількості класів (20 або 200), тому останній класифікаційний шар CaffeNet був замінений на шар з $N + 1$ виходами (з додатковим класом для фону).

Селективний пошук видавав близько 2000 регіонів з різними розмірами і співвідношенням сторін, однак, CaffeNet приймає на вхід тільки зображення фіксованого розміру 227×227 пікселів, тому перед подачею регіонів на вхід мережі їх необхідно модифікувати. Для цього зображення з регіону вставлялося в найменший квадрат, що може охопити його. Вздовж тієї (меншої) сторони, за якою створювалися поля, при цьому, до зображення додавалося декілька контекстних пікселів зображення, частина поля, що залишилися не заповнюється. Отриманий квадрат масштабувався до розміру 227×227 і подався на вхід CaffeNet [36].

Не дивлячись на те, що CNN тренується на розпізнаванні $N + 1$ класів, в результаті вона використовувалась лише для вивільнення фіксованого 4096-розмірного вектору характеристик. Безпосередньо класифікацією зображення займається N лінійних SVM, кожен із яких проводить бінарну класифікацію за своїм типом об'єктів, визначаючи, чи є такий об'єкт на запропонованому зображенні. Схема роботи алгоритму показана на рисунку 3.7.

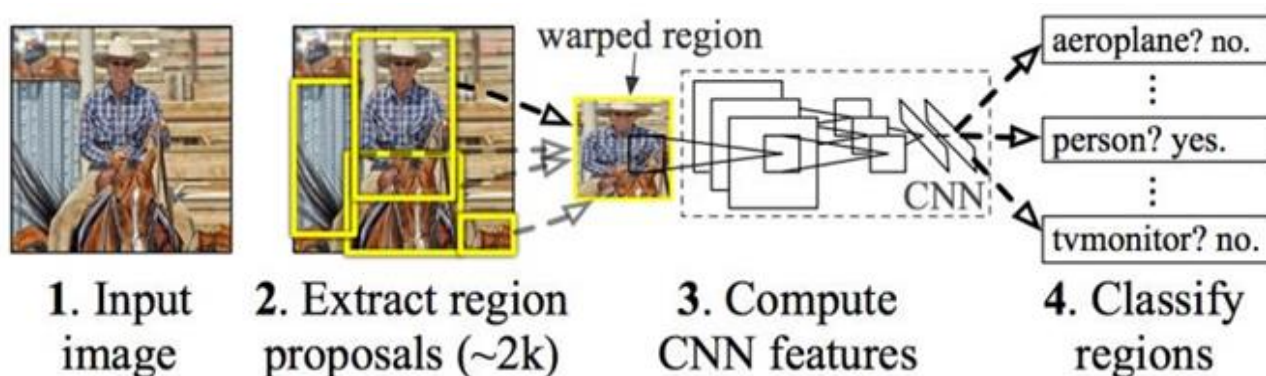


Рисунок 3.7 – Схема роботи алгоритму R-CNN [35]

Процес класифікації в SVM має досить високі показники якості, представляючи собою прості матричні операції. Отримані із CNN вектори характеристик поєднуються по всім регіонам в матрицю розміром 2000×4096 , яка потім множиться на матрицю $4096 \times N$ з вагами SVM.

Отримані за допомогою селективного пошуку регіони тільки можуть містити якісь об'єкти, і не факт, що містять їх цілком. Чи вважати, що регіон містить об'єкт чи ні, визначається за матрицею. Метрика Intersection over Union (IoU) є відношенням площі перетину прямокутного регіону ймовірного об'єкта з прямокутником, що насправді охоплює об'єкт, до площі об'єднання цих прямокутників. Якщо відношення більше за задану граничну величину, регіон- вважається таким, що містить потрібний об'єкт [37].

IoU так само використовувалось для відсіювання надмірної кількості регіонів, що містять певний об'єкт (non-maximum suppression). Якщо IoU деякого регіону відносно регіону, який отримав для того ж об'єкта максимальний результат, був вище деякого значення, перший регіон просто відкидався.

В ході процедури аналізу помилок розробили метод, що дозволяє зменшити помилку виділення рамки, що охоплює об'єкт – bounding-box regression. Після класифікації вмісту регіону, за допомогою лінійної регресії на основі ознак з CNN визначалися чотири параметри – dx , dy , dw , dh . Вони описували, наскільки треба змістити центр рамки регіону по x і y , а також на скільки змінити її ширину і висоту, щоб точніше охоплювати розпізнаний об'єкт [37].

Таким чином, процедуру детектування об'єктів мережею R-CNN можна розділити на наступні кроки:

- а) виділення регіонів-кандидатів за допомогою селективного пошуку;
- б) перетворення регіону в розмір, який приймає CNN CaffeNet;
- в) отримання за допомогою CNN 4096-розмірного вектору ознак;
- г) проведення N бінарних класифікацій кожного вектору ознак за допомогою N лінійних SVM;
- д) лінійна регресія параметрів рамки регіону для більш точного охоплення об'єкта.

3.3.2 Fast R-CNN

Незважаючи на високі результати, продуктивність R-CNN є невисокою, особливо для більш глибоких, ніж CaffeNet мереж (таких як VGG16). Крім того, навчання bounding box regressor і SVM вимагає збереження на диск великої кількості ознак, тому воно дороге з точки зору розміру сховища.[38]

Автори Fast R-CNN запропонували прискорити процес за рахунок таких модифікацій:

а) пропускати через CNN не кожен з 2000 регіонів-кандидатів окремо, а все зображення цілком. Запропоновані регіони потім накладаються на отриману загальну карту ознак;

б) замість незалежного навчання трьох моделей (CNN, SVM, bbox regressor) поєднати всі процедури тренування в одну [38].

Перетворення ознак, які потрапили в різні регіони, до фіксованих розмірів проводилося за допомогою процедури RoIPooling. Вікно регіону шириною w і висотою h поділялося на сітку, що має $H \times W$ осередків розміром $h / H \times w / W$. (Автори документа використовували $W = H = 7$). По кожній такій клітинці проводився Max Pooling для вибору тільки одного значення, даючи таким чином результуючу матрицю ознак $H \times W$.

Бінарні SVM не використовувалися, замість цього отримані ознаки передавалися на повнозв'язний шар, а потім на два паралельних шари: softmax з $K + 1$ виходами (по одному на кожен клас + 1 для фону) і bounding box regressor.

Загальна архітектура мережі показана на рисунку 3.8:

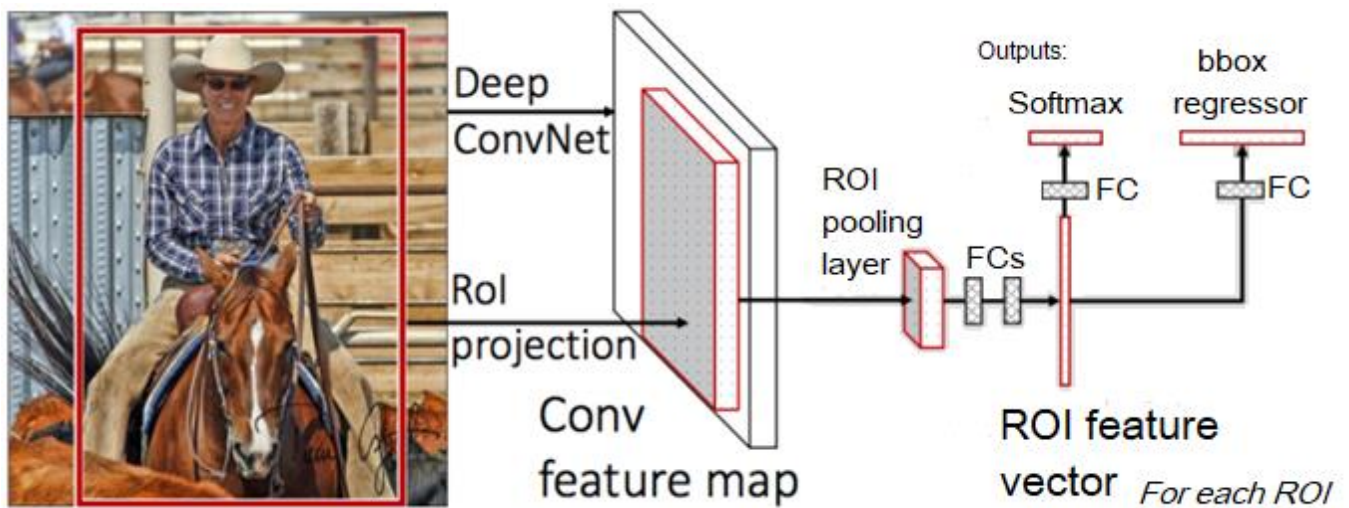


Рисунок 3.8 – Загальна архітектура мережі Fast R-CNN [38]

Для спільного навчання softmax-класифікатора і bbox regressor-a використовувалася об'єднана loss-функція:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda[u \geq 1]L_{loc}(t^u, v) \quad (3.2)$$

де u – клас об'єкта, реально зображеного в регіоні-кандидата;

$L_{cls}(p, u) = -\log p_u$ – логарифм loss для класу u ;

$v = (v_x, v_y, v_w, v_h)$ – реальні зміни рамки регіону для більш точного охоплення об'єкта;

$t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ – передбачені зміни рамки регіону;

L_{loc} – loss-функція між передбаченими і реальними змінами рамки;

$[u \geq 1]$ – індикаторна функція, рівна 1, коли $u \geq 1$, і 0, коли навпаки. Класом $u = 0$ позначається фон (тобто відсутність об'єктів в регіоні);

λ – коефіцієнт, призначений для балансування вкладу обох loss-функцій в загальний результат. У всіх експериментах він однаковий та дорівнює 1 [38].

3.3.3 Faster R-CNN

Після поліпшень, зроблених в Fast R-CNN, вразливим місцем нейромережі виявився механізм генерації регіонів-кандидатів. У 2015 команда з Microsoft Research змогла зробити цей етап значно швидшим. Вони запропонували обчислювати регіони не по початковому зображенню, а знову ж таки по карті ознак, отриманих з CNN. Для цього був доданий модуль під назвою Region Proposal Network (RPN) [39]. Нова архітектура має назву Faster R-CNN та показана на рис. 3.9.

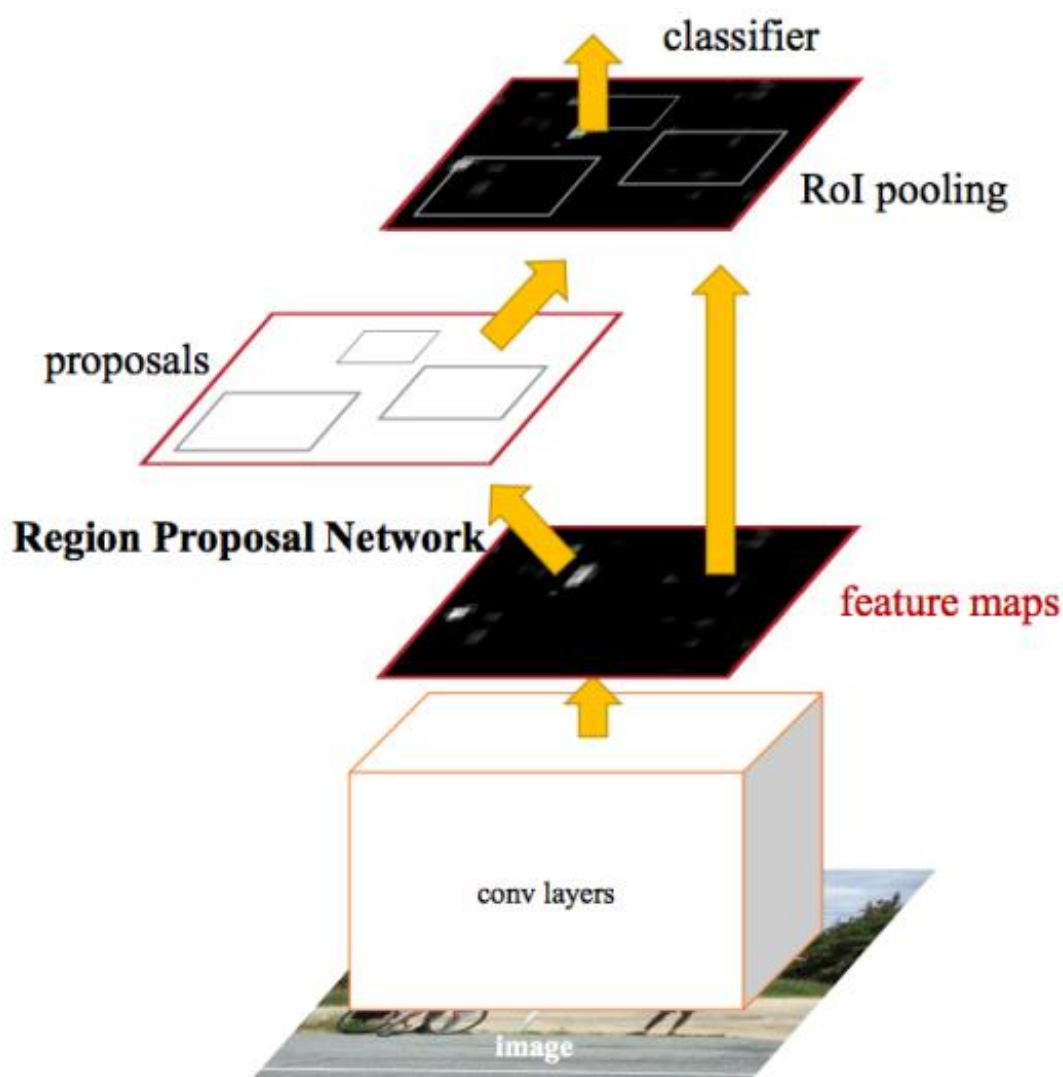


Рисунок 3.9 – Загальна архітектура мережі Faster R-CNN [39]

В рамках RPN за отриманими CNN ознаками використовують малі нейромережі з невеликим (3×3) вікном. Отримані з їх допомогою значення передаються в два паралельних повнозв'язних шари: box-regression layer (reg) і box-classification layer (cls). Виходи цих шарів базуються на так званих anchor-ах: k рамках для кожного положення ковзного вікна, що мають різні розміри і співвідношення сторін. Reg-шар для кожного такого anchor-а видає по 4 координати, що коректують положення охоплюючої рамки; cls-шар видає по два числа - ймовірності того, що рамка містить хоч якийсь об'єкт або що не містить [39]. Робота алгоритму показана на рисунку 3.10.

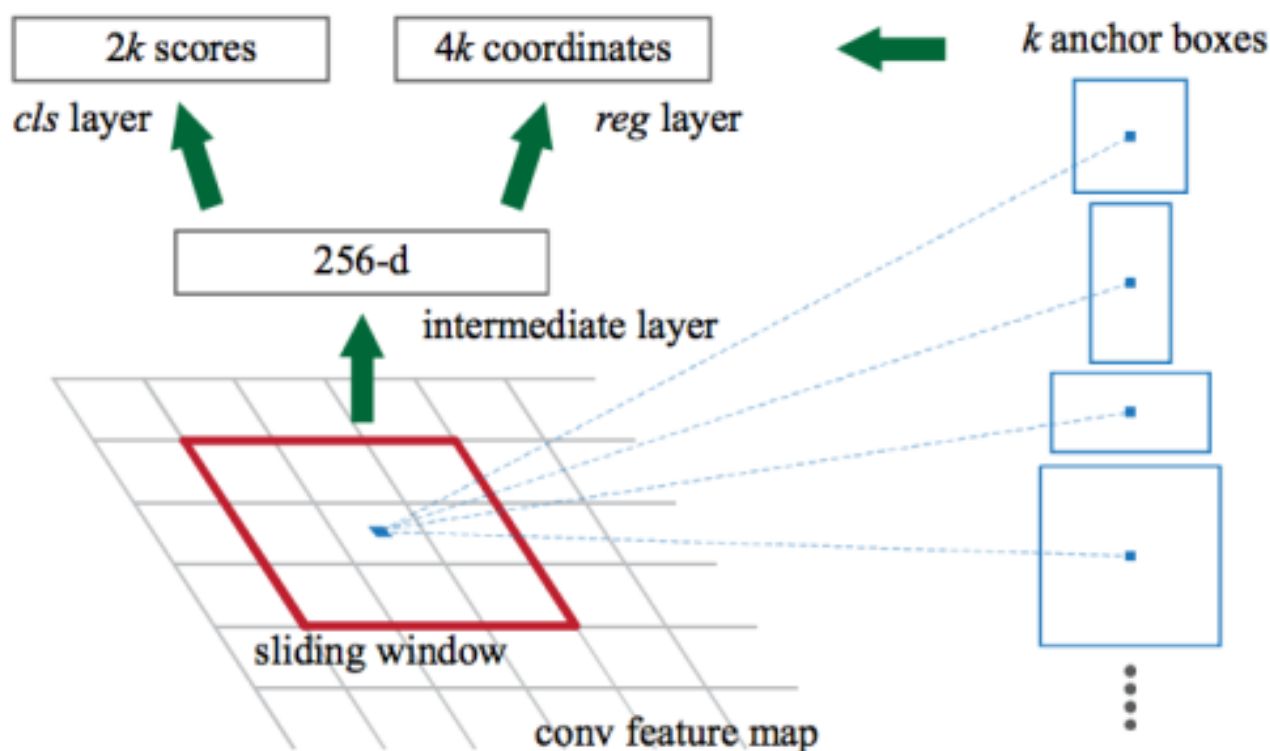


Рисунок 3.10 – Схема алгоритму для побудови карти ознак [39]

Процес навчання reg і cls шарів об'єднаний; loss-функцію вони мають загальну, що представляє собою суму loss-функцій кожного з них, з балансуєчим коефіцієнтом.

Обидва шари RPN видають тільки пропозиції для регіонів-кандидатів. Ті з них, які мають високу ймовірність вмісту будь-якого об'єкта, передаються далі в модуль

детектування об'єктів і уточнення охоплюючої рамки, який як і раніше реалізований як Fast R-CNN.

Для того, щоб розділяти ознаки, одержувані в CNN, між RPN і модулем детектування, процес навчання всієї мережі побудований ітераційно, з використанням декількох кроків:

- а) ініціалізується і навчається на визначення регіонів-кандидатів RPN-частина;
- б) з використанням запропонованих RPN регіонів повторно навчається Fast R-CNN частина;
- в) навчена мережа детектування використовується, щоб ініціалізувати ваги для RPN. Загальні згорткові шари фіксуються і проводиться перенавчання тільки шарів, специфічних для RPN;
- г) з зафіксованими згортковими шарами остаточно переналаштовується Fast R-CNN.

Запропонована схема не є єдиною, і навіть в поточному вигляді вона може бути продовжена подальшими ітераційним кроками, але автори оригінального дослідження проводили експерименти саме після такого навчання.

3.4 Одноетапні детектори

Наскільки перспективним не був би двоетапний детектор, обчислення на другій стадії (тобто, на етапі класифікатора) є досить важкими. Отже, були запропоновані детектори одноетапної дії (SSD), які не покладаються на пропозицію регіону з надією, що вони збільшать швидкість роботи системи. SSD, такі як You Only Look Once (YOLO), розроблені таким чином, що одна мережа передбачає пропозиції регіонів, а також клас пропозицій цих регіонів. Така конструкція економить значну кількість обчислювального часу, дозволяючи виконувати її в 3 рази швидше, ніж найсучасніший Faster R-CNN, досягаючи при цьому більш високої точності.

3.4.1 YOLO

You only look once (YOLO) - це єдина нейронна мережа згортки для прогнозування обмежувальних полів та цільових категорій із повних зображень, яка ділить вхідне зображення на клітини $S \times S$ та прогнозує кілька обмежувальних боксів із ймовірністю їх класу для кожної комірки. Архітектура YOLO складається з вхідного шару, шару згортки, шару об'єднання, повністю зв'язаного шару та вихідного шару. Шар згортки використовується для вилучення ознак зображення, повнозв'язний шар використовується для прогнозу положення зображення та оцінених значень ймовірності цільових категорій, а шар об'єднання відповідає за зменшення пікселів фрагмента [40].

Мережева архітектура YOLO показана на рисунку 3.11.

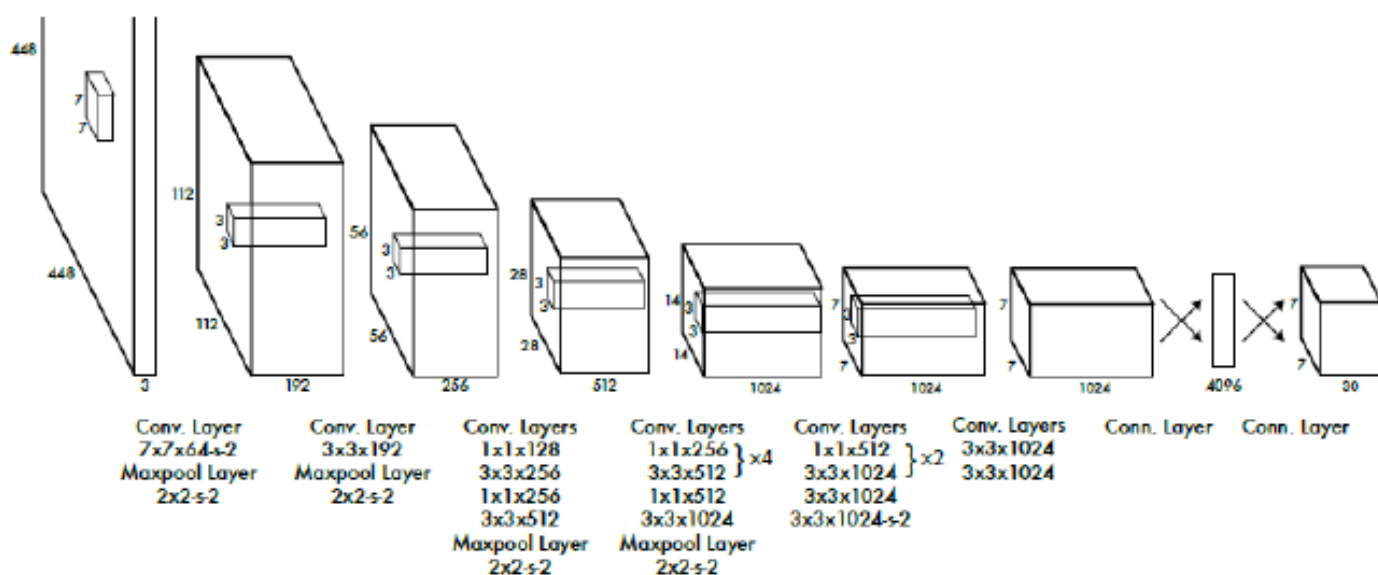


Рисунок 3.12 – Архітектура мережі YOLO [40]

Мережа має 24 згорткових шари, а потім два повністю зв'язаних шари. Чергування згорткових шарів 1×1 зменшує розмір ознак від попередніх шарів. Попередньо тренують згорткові шари для завдання класифікації ImageNet з половиною роздільної здатності (вхідні зображення 224×224), а потім подвоюється роздільна здатність для самого процесу виявлення.

Припустимо, що B – кількість розсувних вікон, що використовуються для кожної комірки для передбачення об'єктів, а C – загальна кількість категорій, тоді розміри вихідного шару – $S * S * (B * 5 + C)$.

Вихідна модель кожної виявленої межі така:

$$T = (x, y, w, h, c), \quad (3.3)$$

де (x, y) являє собою центральні координати обмежувального боксу і (w, h) являє собою висоту і ширину обмежувального вікна виявлення. Наведені вище чотири індекси були нормалізовані щодо ширини та висоти зображення, c - оцінка достовірності, яка відображає значення ймовірності поточного вікна, що містить точність об'єкта виявлення, і формула обчислення наступна:

$$C = P_0 * P_{IOU}, \quad (3.4)$$

де P_0 вказує на ймовірність включення об'єкта виявлення, P_{IOU} вказує співвідношення площі перекриття ковзного вікна та реального виявленого об'єкта.[]

$$P_{IOU} = \frac{Area(BB_i \cap BB_g)}{Area(BB_i \cup BB_g)}, \quad (3.5)$$

де BB_i – це обмежувальний бокс виявленого об'єкта, а BB_g - опорний стандартний бокс на основі навчальної вибірки [40].

3.5 Аналіз структури детекторів

Нижче наведено узагальнену таблицю порівняння двох типів детекторів.

Таблиця 3.1 – Загальна характеристика типів детекторів

Тип	Переваги	Недоліки
Двоетапний	Високий рівень точності при високій ентропії	Низька швидкість та неможливість роботи у режимі реального часу
Одноетапний	Висока швидкість та можливість роботи у режимі реального часу	Більш низький рівень точності

3.6 Висновки

У розділі наведено загальну характеристику методів розпізнавання та класифікації об'єктів. Описано структуру алгоритмів, що використовують різні підходи до виконання процесу розпізнавання. Наведено короткий аналіз по кожному алгоритму.

З наведених даних, можна зробити висновок, що двоетапні методи впроваджуються для підвищення точності та швидкості роботи мережі, тоді як одноетапні методи орієнтуються на загальну швидкість системи, що дозволяє використовувати їх у режимі реального часу. Детальне порівняння роботи різних алгоритмів представлено у наступному розділі.

4 ПОРІВНЯННЯ АЛГОРИТМІВ

4.1 Опис макета

З метою виявлення оптимального алгоритму для створення системи розпізнавання об'єктів постає питання проведення аналізу сучасних алгоритмів розпізнавання та класифікації об'єктів.

Для проведення аналізу алгоритмів було створено програмний застосунок, який написаний з використанням MATLAB. Розроблений застосунок дає змогу провести навчання мережі на одній з доступних вибірок для різних алгоритмів, при цьому отримавши інформацію про якісні характеристики навчання. А також, дає змогу оцінити роботу нейромереж. Лістинг програмного засобу наведено у Додатку Б. Для тренування було використано окремий сервер з віртуальною машиною. Операційна система, що використовувалася – Windows 10 Pro 64 bit. ОЗУ – 20 Гб, ЦП – Intel Xeon CPU E3-1270 v6 3.8 GHz. Для роботи використовувалося середовище MATLAB R2019a з пакетами Deep Learning Toolbox та Computer Vision Toolbox.

У якості тренувальної вибірки було обрано датасет Caltech. Caltech – це вибірка, створена Каліфорнійським технологічним інститутом, що складається приблизно з 10 годин відео розширенням 640x480 та частотою 30 Гц, знятого з транспортного засобу, що здійснює регулярний рух у міському середовищі. Було зроблено анотацію близько 250 000 кадрів (у 137 приблизно хвилинних сегментах) із загальною кількістю 350 000 обмежувальних боксів та 2300 унікальних пішоходів. Анотація включає тимчасову відповідність між обмежувальними боксами та детальною міткою оклюзії.

Обмеження кількості класів до одного, а саме пішоходів зумовлене обмеженістю обчислювальних ресурсів.

Тренування складається з 60 000 кроків для кожного алгоритму. На рисунку 4.1 показано графік зміни показника втрат для різних алгоритмів під час навчання.

На рисунку 4.2 показано приклад вікна застосунка. Для відображення можна обрати необхідну частину датасету та відобразити необхідний кадр. Також є можливість вибору випадкового кадру.

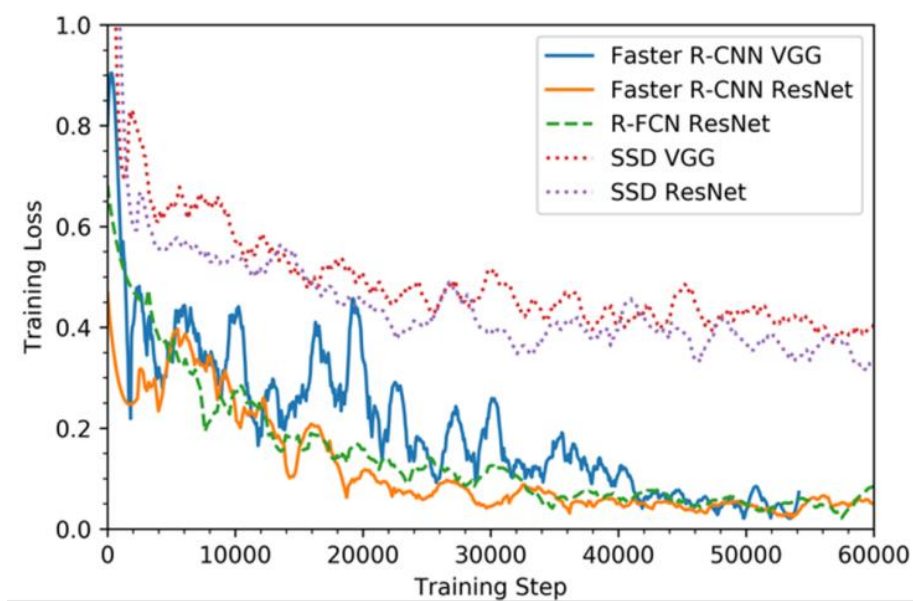


Рисунок 4.1 – Графік показника втрат при навчанні

Для перегляду доступні три види анотації:

- а) результати виявлення мережі;
- б) результати ручної анотації, що використовувалася при тренуванні;
- в) різницю між попередніми результатами.



Рисунок 4.2 – Вікно застосунку

Зміна алгоритму відбувається безпосередньо у програмному коді. Для проведення досліджень було обрано наступні алгоритми: HOG, R-CNN, Fast R-CNN, Faster R-CNN та YOLO.

4.2 Отримані результати

На рисунку 4.3 показано приклад роботи алгоритму HOG.

На рисунку 4.4 показано приклад роботи алгоритму R-CNN.

На рисунку 4.5 показано приклад роботи алгоритму Fast R-CNN.

На рисунку 4.6 показано приклад роботи алгоритму Faster R-CNN.

На рисунку 4.7 показано приклад роботи алгоритму YOLO.

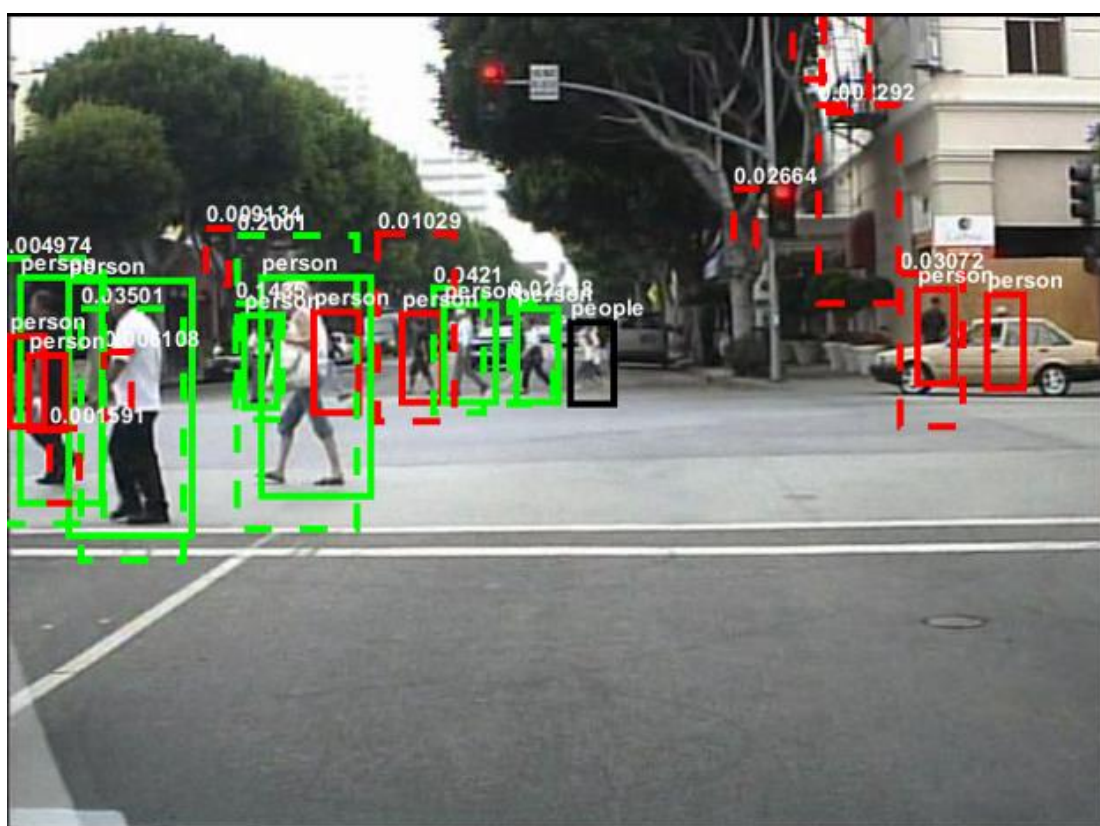


Рисунок 4.3 – Приклад роботи алгоритму HOG

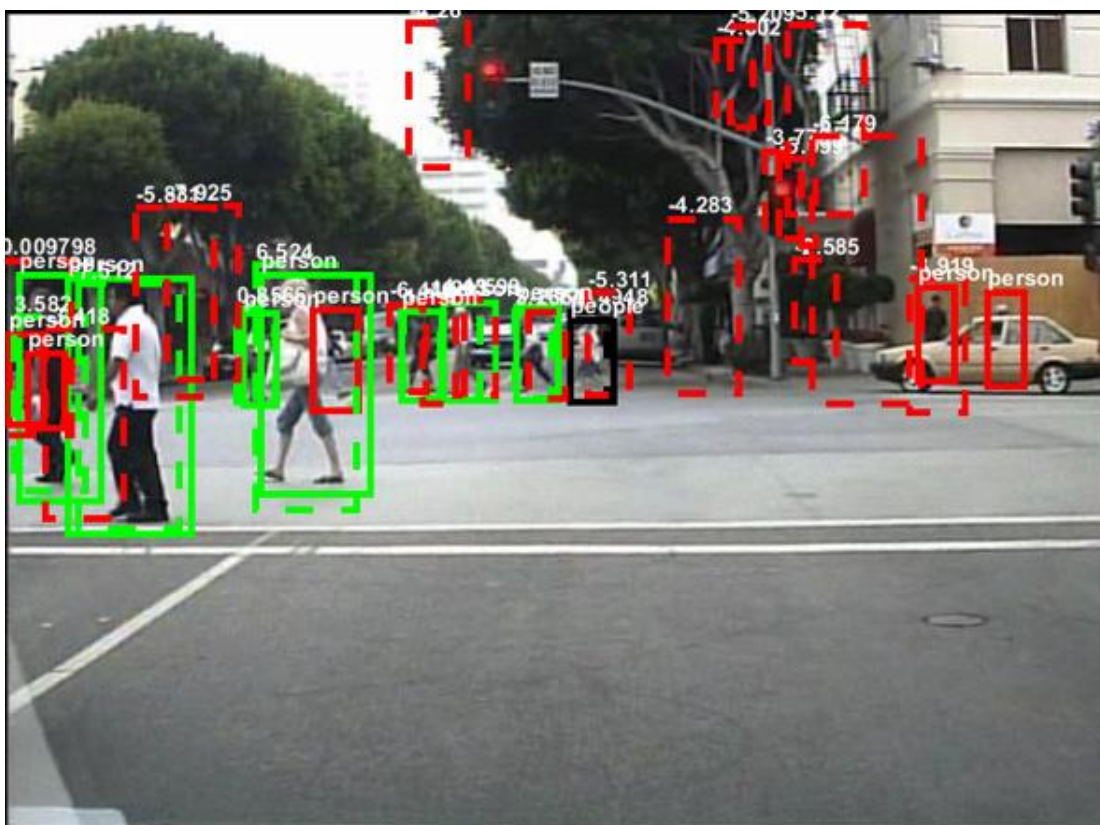


Рисунок 4.4 – Приклад роботи алгоритму R-CNN

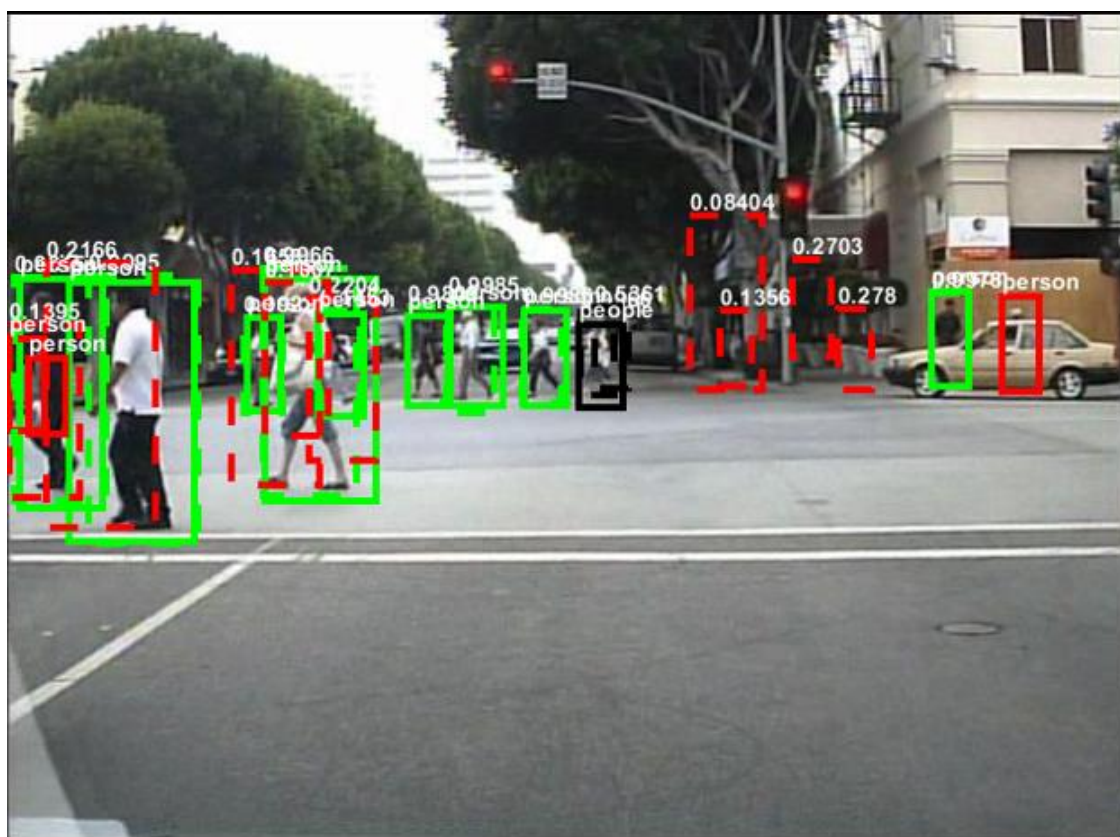


Рисунок 4.5 – Приклад роботи алгоритму Fast R-CNN

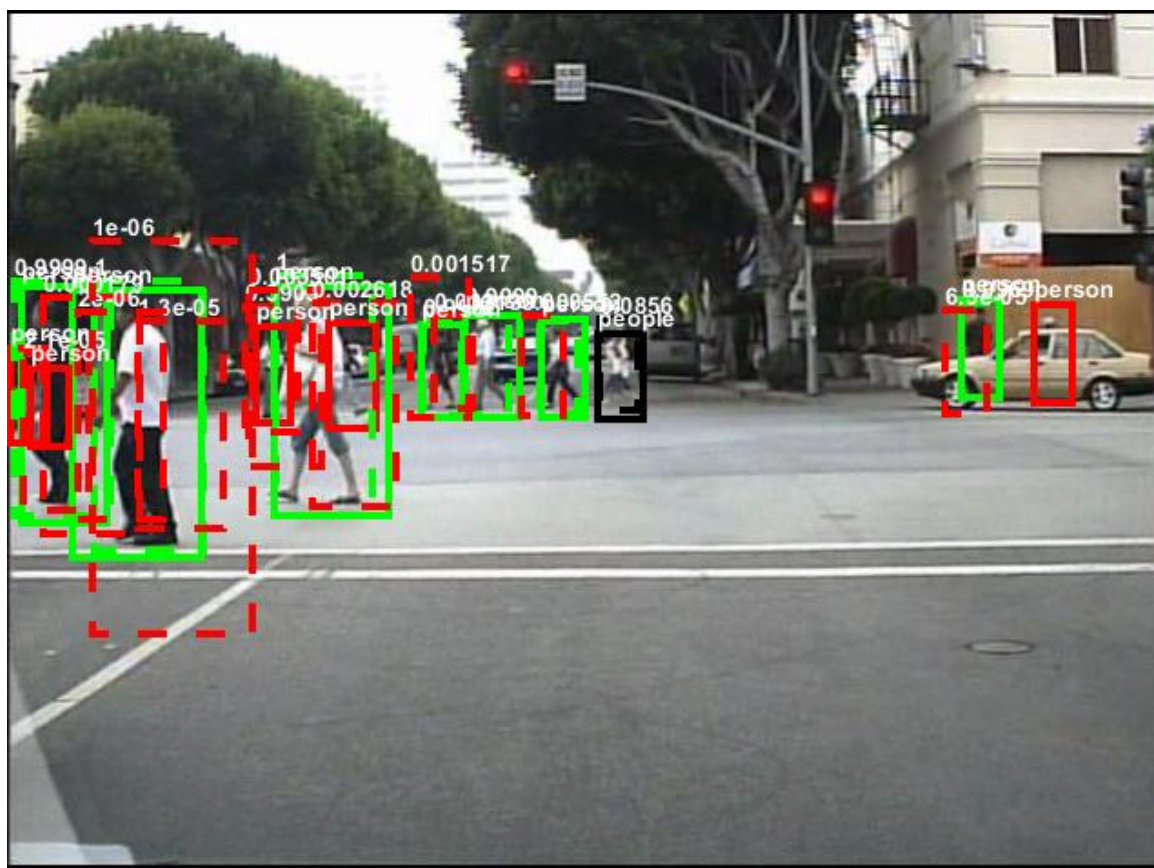


Рисунок 4.6 – Приклад роботи алгоритму Faster R-CNN

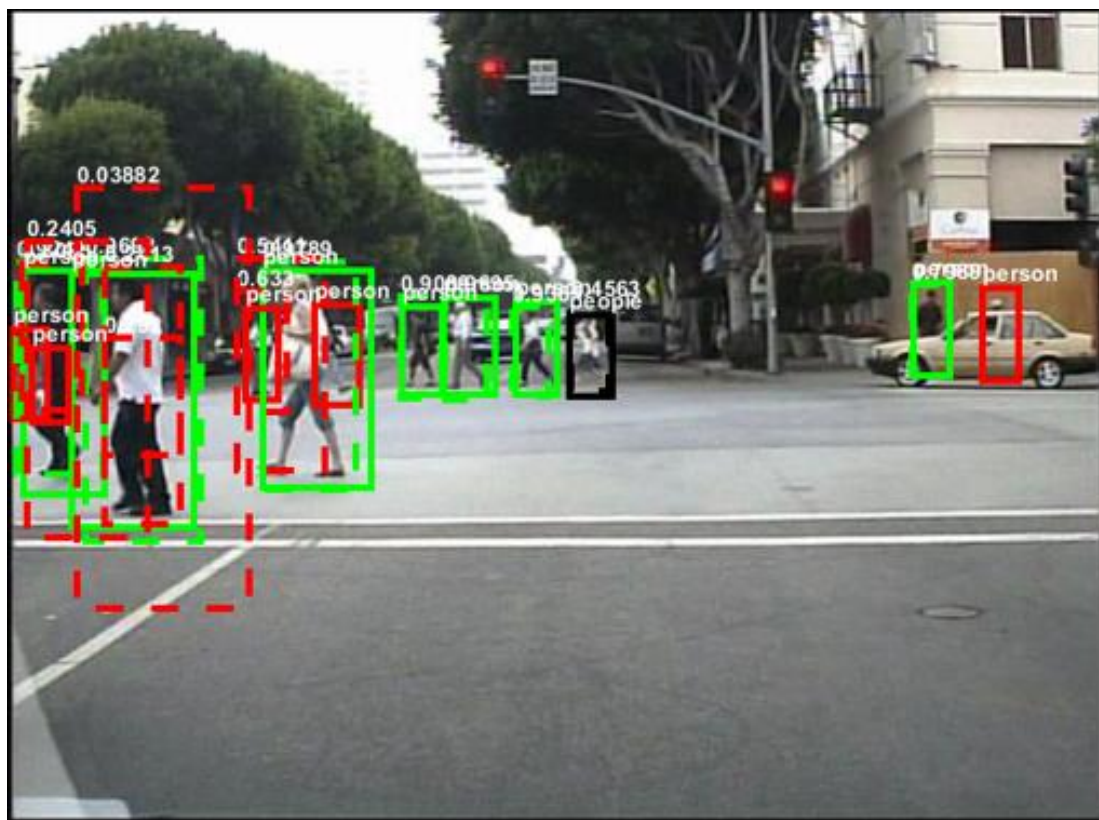


Рисунок 4.7 – Приклад роботи алгоритму YOLO

Для порівняння алгоритмів були складені результуючі таблиці та графіки з показниками якості та швидкості роботи алгоритмів.

На рисунку 4.8 представлено графік співвідношення хибно позитивних знайдених об'єктів до загальної кількості об'єктів. Хибно позитивний результат означає, що мережею було знайдено об'єкт, якого насправді там немає. У таблиці 4.1 показано результуючі показники співвідношення хибно позитивних знайдених об'єктів до загальної кількості об'єктів.

Як видно з графіку, найгірший результат показав алгоритм HOG, а найкращий Faster R-CNN, при цьому показники YOLO близькі до Faster R-CNN.

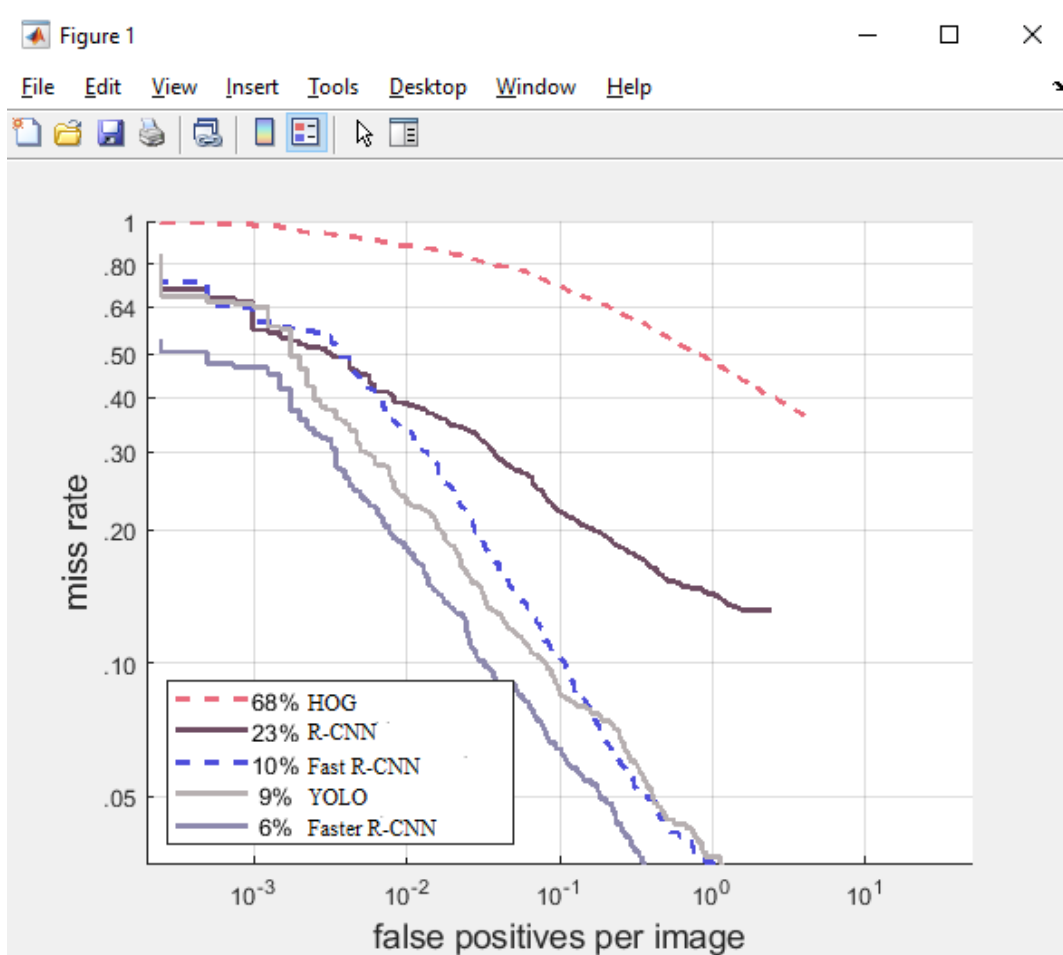


Рисунок 4.8 – Відношення кількості хибно позитивних спрацювань до загальної кількості об'єктів

Таблиця 4.1 – Результуючі показники співвідношення хибно позитивних знайдених об’єктів до загальної кількості об’єктів

Метод	Miss rate
HOG	68%
R-CNN	23%
Fast R-CNN	10%
YOLO	9%
Faster R-CNN	6%

Для оцінки точності роботи використаємо такий показник як mAP – середня точність. Такий показник використовується у глибинному навчанні оцінки того, наскільки результати пошуку задовольняють запит користувача. Алгоритм розрахунку наступний – беремо всю множину результатів розпізнавання, зазвичай не всі вони правильні, тому обчислюємо точність на кожному правильно поверненому зображенні, а потім беремо середнє значення. Наприклад, якщо повернутий результат дорівнює 1, 0, 0, 1, 1, 1, де 1 – зображення пішохода, а 0 ні, точність на кожному правильному зображенні дорівнює відповідно – 1/1, 0, 0, 2/4, 3/5, 4/6, сум точності – 83/30, а середня точність – (сума точності) / (повне правильне зображення) = 83/120.

У таблиці 4.2 приведено середні точності для кожного з досліджуваних алгоритмів.

Таблиця 4.2 – Середня точність алгоритмів

Метод	mAP
HOG	14.2
R-CNN	29.7
Fast R-CNN	31.9

Продовження таблиці 4.2

Метод	mAP
YOLO	43.0
Faster R-CNN	46.8

Як видно з даних таблиці 4.2, фінальні результати є досить невеликими, а Faster R-CNN та YOLO показують кращі результати. Але такий рівень точності говорить про те, що такі алгоритми не можна використовувати без додаткових програмних або апаратних засобів, що дадуть змогу підвищити показники якості.

Наступним і не менш важливим показником для оцінки того, чи можна використовувати дані алгоритми у реальних системах є кількість кадрів, оброблюваних за секунду (fps), або кількість часу, що витрачається на обробку одного кадру. У таблиці 4.3 наведено показники кадрів за секунду для аналізованих алгоритмів.

Таблиця 4.3 – Кількість кадрів за секунду

Метод	fps
HOG	0.1
R-CNN	0.2
Fast R-CNN	0.5
YOLO	45.0
Faster R-CNN	1.2

Для більш наочного відображення отриманих результатів побудуємо діаграму часу, потрібного на обробку одного кадру. Діаграму показано на рисунку 4.8.

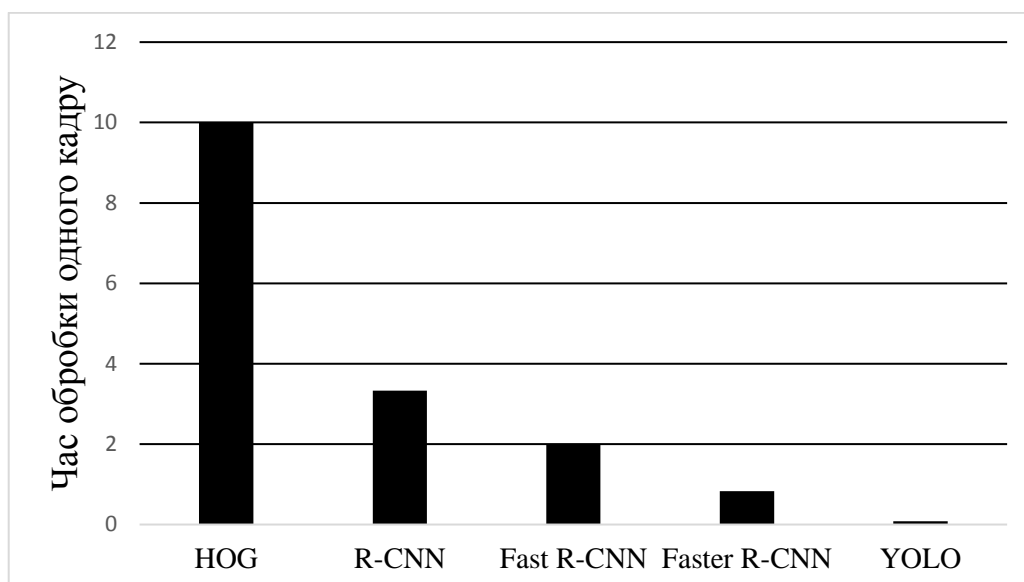


Рисунок 4.8 – Час обробки одного кадру

4.3 Аналіз отриманих результатів

Як видно з отриманих результатів, тільки алгоритм YOLO може бути використаний для систем розпізнавання в реальному часі. Інші алгоритми можна з впевненістю використовувати для офлайн детекторів. Однак, враховуючи точність розпізнавання YOLO, необхідно виявити найкращі шляхи підвищення точності. Використання додаткових алгоритмів обробки збільшує час обробки одного кадру, а також потребує додаткових обчислювальних засобів, крім того, такий підхід не дасть змоги значно підвищити показники якості. Отже, треба використовувати апаратні рішення. У наступному розділі буде розглянуто можливі шляхи підвищення точності зі збереженням швидкодії існуючого алгоритму.

4.4 Висновки

У розділі описано розробку застосунку для порівняння різних типів алгоритмів в однакових умовах для виявлення найкращого, з точки зору компромісу швидкодії та точності. Програмно реалізовано запропоновані алгоритми та проведено аналіз показників якості та швидкодії їх роботи.

5 КОНЦЕПЦІЯ СИСТЕМИ

5.1 Структура системи

Аналізуючи результати порівняння алгоритмів розпізнавання, отримані в попередньому розділі, було вирішено використовувати алгоритм YOLO, проте, для підвищення точності та якості розпізнавання поєднати його з лідаром.

Обробка об'ємних масивів точок стає все більш важливою для розвитку автономних транспортних засобів, це відбувається завдяки стрімкому розвитку лідарів за останні декілька років. Сучасні лідари здатні працювати з великими об'ємними моделями у режимі реального часу, при цьому їх ціна стрімко зменшилася з майже 100 000 доларів до 100 доларів. Перевагою використання лідару є можливість отримання положення різних об'єктів у просторі з дуже високою точністю. А також, можливість оцінювати напрям руху об'єктів. Порівняно із плоскими зображеннями, розподілення точок у просторі є більш рідким з різним ступенем розподілення по всій площі. Ці точки є не упорядкованими, вони взаємодіють локально і в основному їх не можна аналізувати ізольовано. Обробка точкових хмар завжди повинна бути інваріантною до основних перетворень.

Загалом, виявлення та класифікація об'єктів на основі глибинного навчання є добре вивчено сферою і має безліч рішень для двовимірної моделі простору. Основна увага в дослідженні була в основному компромісом між точністю та ефективністю. Що стосується систем автономного управління, ефективність у цій галузі набагато важливіша. Тому найкращі детектори об'єктів використовують мережі регіональних пропозицій (RPN) або подібний RPN-підхід із сіткою. Ці мережі надзвичайно ефективні, точні та навіть здатні працювати на спеціальному апаратному або вбудованому пристрої. Виявлення об'єктів на точкових хмарах все ще не має такого розповсюдження, але є більш важливим завданням. Реалізація 3D моделі передбачає можливість створення об'ємних обмежувальних боксів. В даний час існує в основному три різні підходи з використанням глибинного навчання:

а) обробка точкової хмари за допомогою багат шарового перцептрону;

б) переведення точкових хмар у вокселі чи стеки зображень за допомогою згорткових нейронних мереж (CNN);

в) комбіновані підходи до синтезу [41];

Поєднання лідару та YOLO утворює ефективну модель, яка працює безпосередньо на RGB-картах, що базуються на поєднанні інформації, отриманої з камери та лідару, щоб оцінити та локалізувати точні тривимірні обмежувальні поля. На рисунку 5.1 показано принцип роботи системи. У верхній частині рисунку зображено вид згори на основі хмари точок лідару Yelodyne HDL64 [42]. У нижній частині зображено кадр з камери з нанесеними об'ємними боксами виявлених та класифікованих об'єктів.

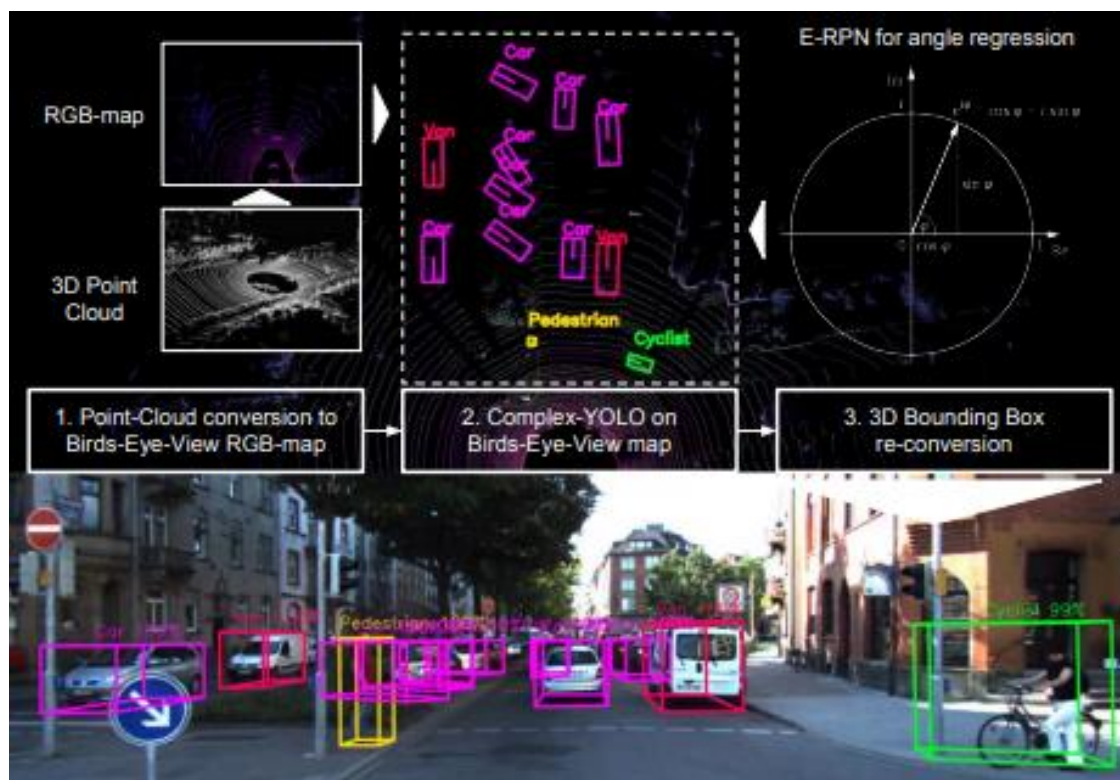


Рисунок 5.1 – Принцип роботи системи [42]

5.2 Використаний алгоритм

Об'ємний алгоритм YOLO базується на специфічному підході E-RPN, який оцінює орієнтацію об'єктів, що описуються уявною та реальною частиною для

кожного боксу. Ідея полягає у тому, щоб мати закритий математичний простір без виявлених ознак для точного узагальнення кутів. Модель здатна передбачати точні тривимірні поля з локалізацією та точною класифікацією об'єктів у режимі реального часу, навіть якщо об'єкт описується кількома точками (наприклад, пішохід). Крім того, алгоритм здатний розпізнавати всі вісім класів датасету KITTI [43]. Оцінивши модель за еталонним тестом KITTI, точність досягла досить високих результатів і займає одну з лідируючих позицій серед доступних у вільному доступі моделей.

3D-хмара точок одного кадру, що отримана за допомогою лідару Velodyne HDL64 перетворюється в єдину RGB-карту з видом з висоти пташиного польоту. що охоплює площу 80 м x 40 м (рисунок 5.2).



Рисунок 5.2 – RGB-карта площі виявлення

RGB-карта описується висотою, інтенсивністю та щільністю. Розмір карти сітки визначається з $n = 1024$ та $m = 512$. Отже, була спроектована та дискретизована 3D хмара точок у двовимірну сітку з роздільною здатністю приблизно $g = 8$ см. Порівняно з MV3D, було зменшено розмір комірки для досягнення менших помилок квантування, що супроводжується більш високою вхідною роздільною здатністю. Через високу ефективність немає потреби використовувати більше однієї карти. Отже, всі три канали ознак $(z_r, z_g, z_b, \text{ де } z_{r,g,b} \in \mathbb{R}^{m \times n})$ обчислюються для всієї хмари точок $P \in \mathbb{R}^3$ всередині області покриття Ω . Хмара точок P описується за допомогою наступної формули [44]:

$$P_{\Omega} = \{P = [x, y, z]^T | x \in [0, 40 \text{ м}], y \in [-40 \text{ м}, 40 \text{ м}], z \in [-2 \text{ м}, 1.25 \text{ м}] \} \quad (5.1)$$

Вибираємо $\in [-2 \text{ м}, 1.25 \text{ м}]$, розглядаючи положення лідару з координатою по осі $z - 1,73 \text{ м}$, щоб охопити площу над землею приблизно до 3м висоти, вважаючи найвищими об'єктами вантажні автомобілі. За допомогою калібрування визначаємо функцію відображення $S_j = f_{PS}(P_{\Omega_i}, g)$ при $S \in \mathbb{R}^{m*n}$ показуючи кожну точку з індексом i у спеціальну комірку сітки S_j нашої RGB-карти. Набір точок описує всі точки, відображені у конкретній комірці сітки[44]:

$$P_{\Omega_{i \rightarrow j}} = \{P_{\Omega_i} = [x, y, z]^T \mid S_j = f_{PS}(P_{\Omega_i}, g)\}. \quad (5.2)$$

Отже, можемо розрахувати канал кожного пікселя, враховуючи інтенсивність точок лідару як I

$$\begin{aligned} z_g &= \max(P_{\Omega_{i \rightarrow j}} * [0, 0, 1]^T), \\ z_b &= \max(I(P_{\Omega_{i \rightarrow j}})), \\ z_r &= \min(1.0, \log(N + 1) / 64), N = |P_{\Omega_{i \rightarrow j}}|, \end{aligned} \quad (5.3)$$

де, N описує кількість точок відображених з P_{Ω_i} на S_j ,
а g параметр розміру комірки сітки [44].

Отже, z_g представляє максимальну висоту, z_b – максимальну інтенсивність, z_r – нормалізовану щільність усіх точок, відображених у S_j . Алгоритм швидкої та точної оцінки 3D поля на хмарі точок наступний – RGB-карта подається в CNN. Сітка E-RPN починає працювати одночасно на останній карті ознак та пропонує п'ять боксів для кожної комірки сітки. Кожне запропоноване поле складається з параметрів регресії та оцінки об'єкта p із загальною ймовірністю p_0 та n оцінками класів (Рисунок 5.3).

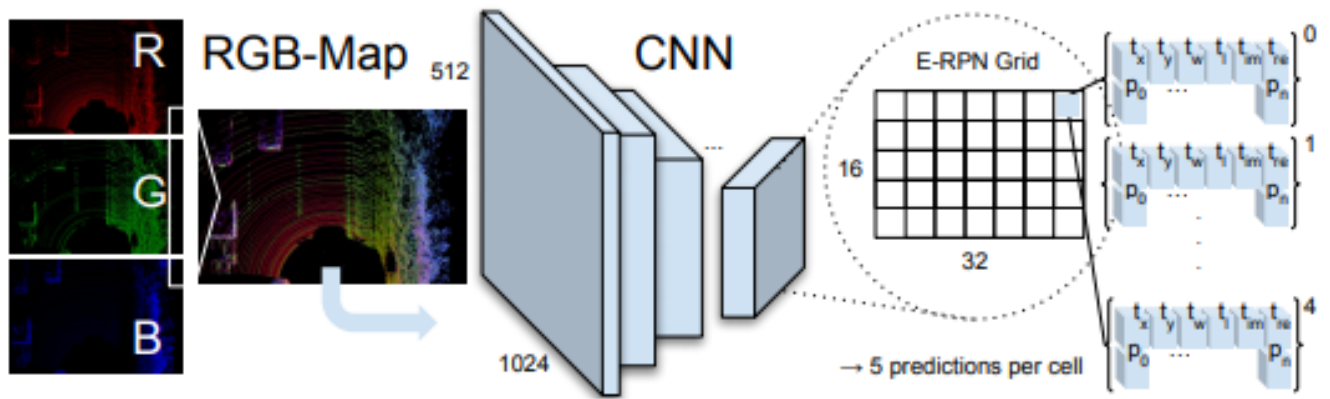


Рисунок 5.3 – Побудова карти зображення [44]

Мережа YOLO приймає на вхід RGB-карту. В моделі використовується YOLOv3 [45] архітектура CNN, розширена за допомогою складної кутової регресії та E-RPN для виявлення точних позицій 3D-об'єктів.

Euler-Region-Proposal. E-RPN аналізує позицію у тривимірному просторі, розміри об'єкта (ширина і довжина), а також ймовірність приналежності об'єкту до певного класу. Аби отримати правильне положення, було змінено один зі стандартних підходів Grid-RPN шляхом додавання до координат положення ще й кута $\arg(|z|e^{ib_\varphi})$:

$$\begin{aligned}
 b_x &= \sigma(t_x) + c_x, \\
 b_y &= \sigma(t_y) + c_y, \\
 b_w &= p_w e^{t_w}, \\
 b_l &= p_l e^{t_l}, \\
 b_\varphi &= \arg(|z|e^{ib_\varphi}) = \arctan_2(t_{l_m}, t_{r_e}).
 \end{aligned} \tag{5.4}$$

За допомогою цього розширення, E-RPN оцінює точну орієнтацію об'єктів на основі уявної та реальної частини, безпосередньо вбудованої в мережу. Для кожної комірки сітки (32x16) прогнозується п'ять об'єктів, включаючи оцінку ймовірності та оцінку класу приналежності, в результаті чого отримаємо 75 ознак.

комплексних чисел, які мають закритий математичний простір для порівняння кутів [46].

$$\mathcal{L}_{Euler} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} |(e^{ib\varphi} - e^{i\hat{b}\varphi})^2| = \quad (5.5)$$

$$= \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(t_{im} - \hat{t}_{im})^2 + (t_{re} - \hat{t}_{re})^2], \quad (5.6)$$

де λ_{coord} – коефіцієнт масштабування для забезпечення стабільної конвергенції на ранніх фазах;

1_{ij}^{obj} – позначає, що прогноз j -го обмежувального поля в комірці i має більший перетин порівняно з істинним положенням об'єкта.

Загальний вигляд взаємодії камери та лідару показано на рисунку 5.5.

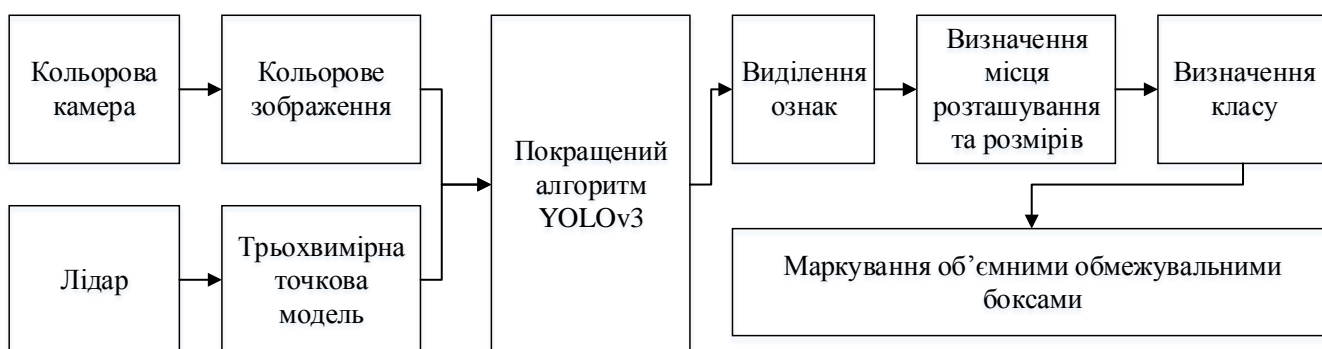


Рисунок 5.5 – Взаємодія камери та лідару

5.5 Реалізація та отримані результати

Використовуючи мову програмування Python, претреновану мережу KittiYOLODataset, а також бібліотеки для комп'ютерного зору було реалізовано алгоритм розпізнавання об'єктів. Для прикладів було обрано доступні кадри та відеоролики із згенерованими з використанням лідару Yelodyne HDL64 об'ємними просторовими моделями. Фрагменти лістингу програмної реалізації представлено у

Додатку В. Реалізація заснована на модифікованій версії нейронної мережі Darknet [28]. Спочатку, застосовується попередня обробка для створення RGB-карти із даних, отриманих від лідару.

Використовуючи загальні принципи машинного навчання, навчальну вибірку було розділено на дві множини – 85% для тренування та 15% для валідації отриманої мережі. Розподіл класів приблизно наступний – 75% автомобілів, близько 4% велосипедистів і близько 15% пішоходів. Результати тренування показали добрі результати для валідаційних даних з різними умовами – місто, траси, умови поганої освітленості тощо. На рисунках 5.6 – 5.7 показано приклади роботи системи.



Рисунок 5.6 – Приклад роботи системи

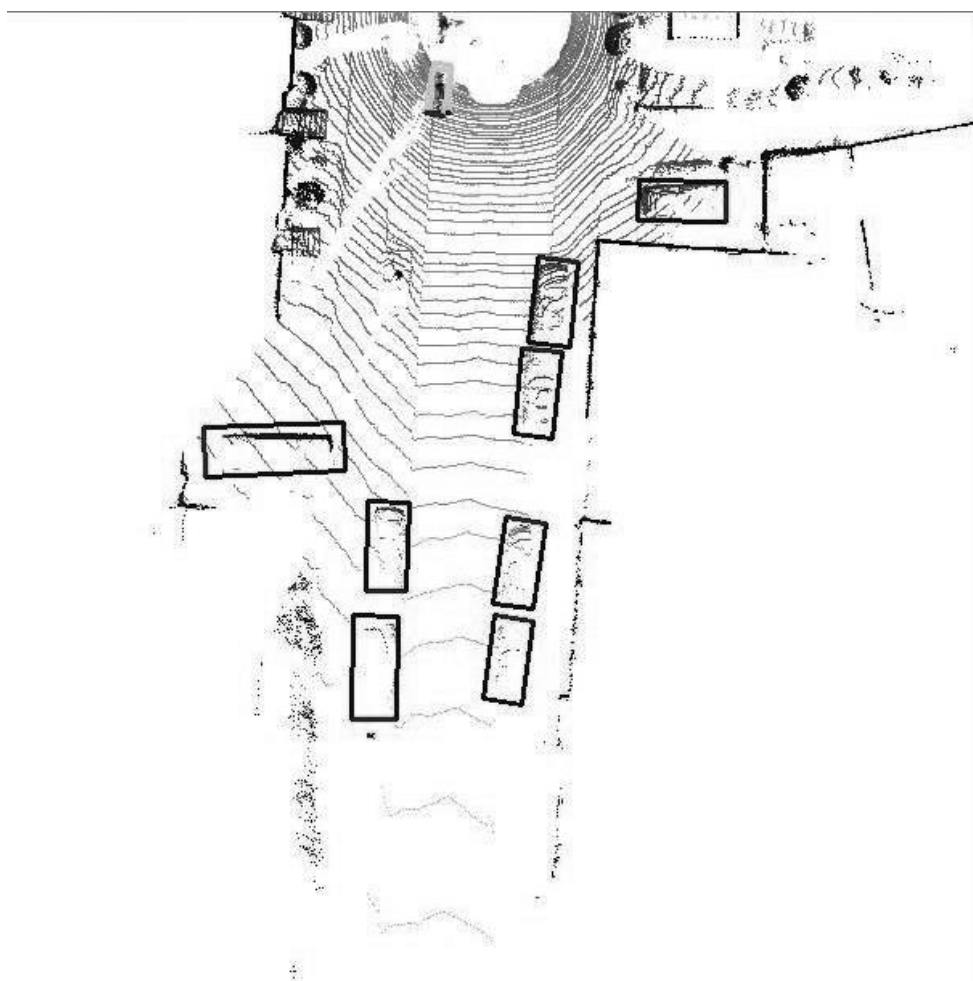


Рисунок 5.7 – Приклад роботи системи

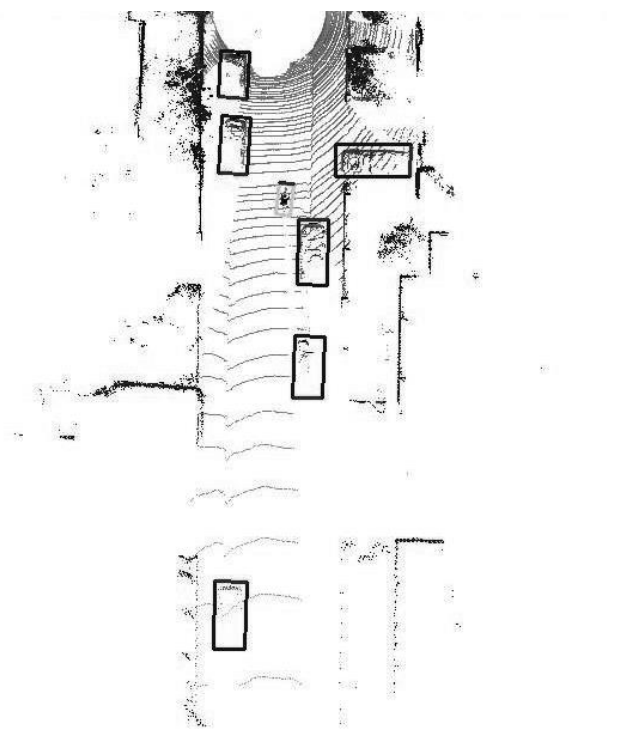


Рисунок 5.8 – Приклад роботи системи

Отримані показники якості роботи системи зведені у таблиці 5.1.

Таблиця 5.1 – Показники якості роботи

Показник	Значення
Кількість кадрів за секунду(FPS)	50.4
Середня точність для класу машин	67.72 %
Середня точність для класу пішоходів	42.79 %
Середня точність для класу велосипедистів	58.32 %

Аналізуючи отримані результати, можна з впевненістю стверджувати, що запропонована модель, що поєднує використання кольорової камери та лідару є оптимальною з розрахунку вартості та отримуваної точності та швидкості класифікації. Таку систему можливо використовуватися для реальних транспортних засобів. Використання лідару відкриває великі можливості, пов'язані з визначення положення як самого транспортного засобу, так і об'єктів навколо. Камера дозволяє більш точно відображати досліджуване середовище, а також надає можливість використовувати отримані дані для пошуку таких речей як лінії розмітки, дорожні знаки, вказівники та визначати сигнали світлофора. Таке поєднання дозволяє створювати досить ефективні системи автономного керування.

5.6 Висновки

У розділі описано структуру побудови запропонованої системи, описано принципи роботи її окремих частин та алгоритмів. Також, описано програмну реалізацію системи та проведено аналіз отриманих результатів. Зроблено висновок щодо можливості використання системи у реальному житті.

6 РОЗРОБЛЕННЯ СХЕМИ ЕЛЕКТРИЧНОЇ СТРУКТУРНОЇ

6.1 Опис

Для опису системи загалом, було розроблено схему електричну структурну. Вона представлена на рисунку 6.1.

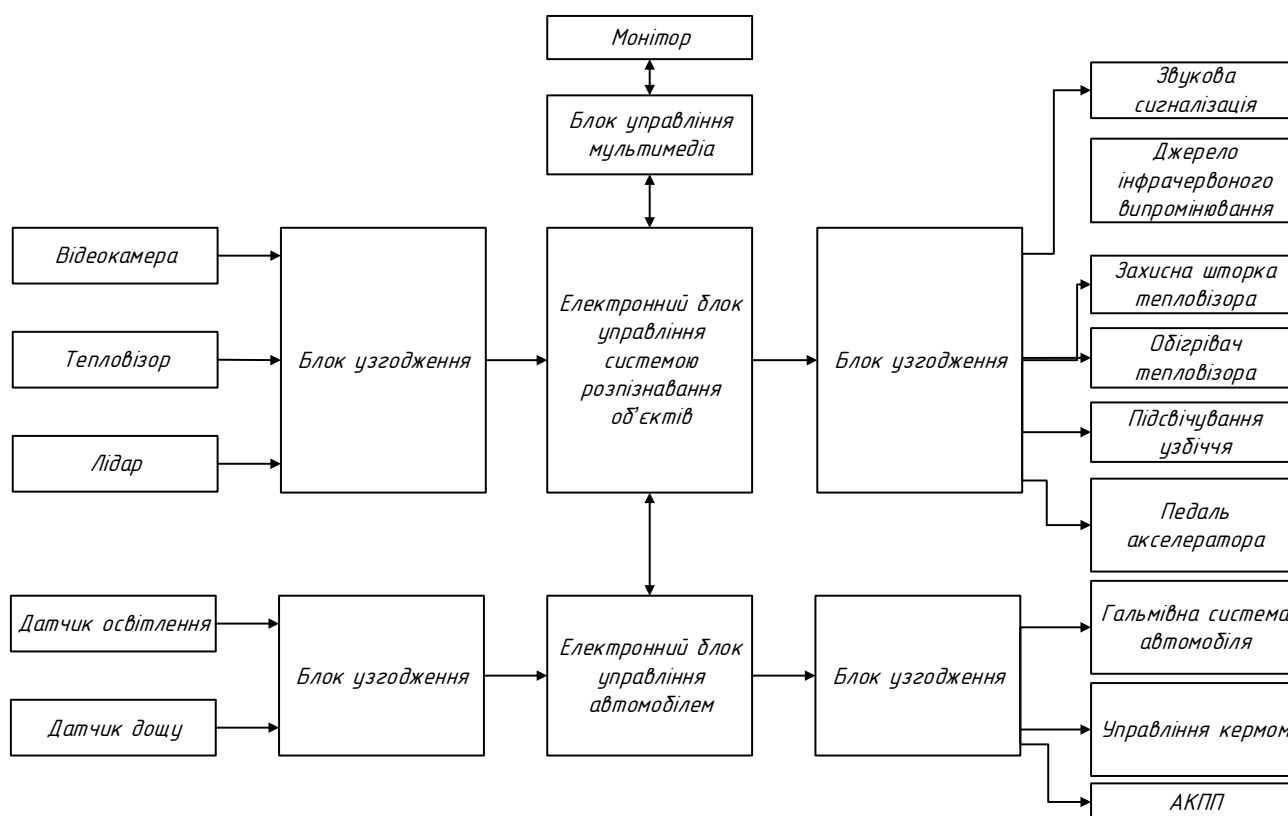


Рисунок 6.1 – Структурна схема системи

В попередніх розділах було детально описано структуру підсистеми розпізнавання об'єктів, однак, для більшої деталізації розроблялася структурна схеми системи загалом.

Структурні схеми описують основні функціональні частини систем, їх структуру та взаємозв'язок між ними. Структурні схеми не описують функціональних особливостей елементів, що входять до складу системи та не дають уявлення про способи зв'язки між ними.

Розроблена схема представляє структуру системи автономного керування у більш загальному вигляді.

Усі представлені елементи можна умовно розділити на три групи.

До першої групи входять давачі інформації, тобто різні датчики, сканери тощо. На представленій схемі до давачів інформації можна віднести кольорову камеру, тепловізор, лідар. Перераховані елементи входять до складу підсистеми розпізнавання об'єктів як основі засоби отримання інформації. Також, до давачів інформації можна віднести датчики освітленості та дощу. Ці елементи входять до підсистеми засобів допомоги водію та систем комфорту. На основі отриманої з їх допомогою інформації електронний блок управління приймає рішення про ввімкнення або вимкнення та зміну режимів роботи денних ходових вогнів, ближнього та дальнього світла головних фар, склоочищувачів, різних систем підігріву скла, дзеркал тощо.

Отримана первинна інформація потрапляє далі до так званих блоків узгодження, які допомагають привести перетворити отримані дані у вигляд, з яким можуть працювати блоки управління. Саме блоки узгодження слугують каналами зв'язку між різними системами.

Друга група пристроїв представлена блоками управління. Зазвичай, блоки управління виконані у вигляді програмованих інтегральних мікросхем або мікроконтролерів, що розміщуються на друкованих печатних платах. Блоки управління приймає на вхід інформацію від давачів та основі проведеного аналізу приймає рішення про виконання певних дій. В сучасних транспортних засобах блоки управління розташовуються в кожній системі, яка існує в авто. Такий децентралізований підхід дає змогу підтримувати робочий стан всієї системи при виході із ладу якоїсь меншої підсистеми. На сьогоднішній день існує близько 100 різних підсистем, що можуть використовуватися в транспортних засобах.

На представленій схемі показано електронний блок управління системою розпізнавання об'єктів. Його структуру показано узагальнено, адже, до його складу також входить підсистеми нічного бачення та інші. Так само узагальнено показано структуру блоку управління автомобілем, до складу якої входять система гальмування, система управління розподілом палива, різні стабілізуючі системи тощо.

Приймаючи дані на вхід, електронний блок у разі потреби генерує управляючі сигнали призначені для виконавчих пристроїв.

Третя група – група виконавчих пристроїв. До них відносяться різні пристрої, що забезпечують роботу системи. Умовно можна сказати, що це вихід системи. До складу цієї групи можуть входити різні двигуни, ліхтарі, обігрівачі тощо.

У розроблюваній системі до цієї групи входять такі елементи: звукова сигналізація, інфрачервоний випромінювач, обігрівач тепловізора, джерело світла для підсвічування узбіччя, електронна педаль акселератора, гальмівна система, автоматична коробка переключення передач.

Для зв'язку між блоками управління та виконавчими пристроями також постає необхідність використання блоків узгодження, адже усі виконавчі пристрої мають різну будову і принципи роботи, а отже і потребують різних видів сигналів для управління.

Так як всі системи так, чи інакше пов'язані між собою, то потрібно забезпечити взаємозв'язок між електронними блоками управління. Такий підхід дозволяє використовувати отриману від давачів інформацію не тільки в межах однієї ізольованої підсистеми, але і в будь-якій частині системи.

Також, різні системи мають різний ступінь пріоритету, тому існує потреба використання або пріоритетних методів передачі інформації, або використання окремих каналів зв'язку між такими системами. До таких систем в першу чергу відносяться системи, що відповідають за безпеку рух – це всі підсистеми, що пов'язані з двигуном, гальмівною системою, рульовим управлінням та аварійні системи(подушки безпеки, преднатягачі пасків безпеки тощо). До систем з нижчим пріоритетом можна віднести системи комфорту, управління мультимедією тощо.

6.2 Висновки

У розділі розглянуто побудову схеми електричної структурної для розроблюваної системи. Наведено перелік елементів та принципи їх поєднання у систему.

7 РОЗРОБЛЕННЯ СХЕМИ ЕЛЕКТРИЧНОЇ ФУНКЦІОНАЛЬНОЇ

7.1 Опис

На рисунку 7.1 представлено функціональну схему частини системи автономного управління, яка відноситься до підсистеми розпізнавання об'єктів та їх трекінгу.

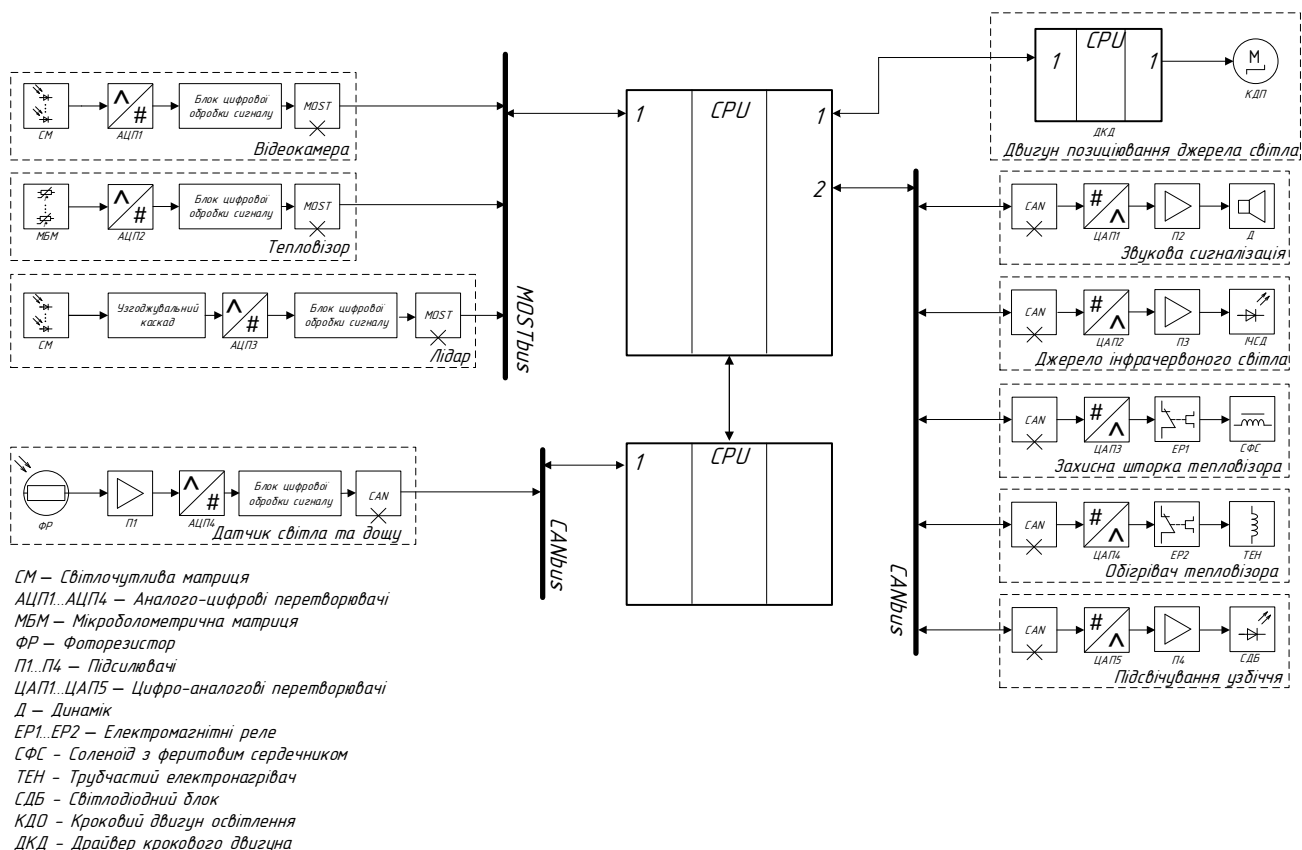


Рисунок 7.1 – Функціональна схеми системи

Загалом, функціональна схема описує функціональні особливості та принципи побудови і роботи окремих елементів системи.

Функціональна структура кольорової камери та лідару дуже схожі, обидва прилади складаються з чутливого сенсору, сингали з яких надходять до аналого-цифрового перетворювача, на виході ми маємо дані у цифровому вигляді і після внутрішньої попередньої обробки ці дані надходять до мікроконтролера для подальшої роботи з ними. Головною функціональною відмінністю є тип сенсорів.

Камера використовує матрицю на основі технології КМОН (комплементарна структура метал-окисел-напівпровідник), де за світлосприйняття відповідає фотодіод, а польові транзистори використовуються для збереження заряду діода та дають можливість сприйняття інформації з одного конкретного фотодіоду. Такий тип використаної матриці значно збільшує енергоефективність, а також за рахунок єдності технологій дозволяє об'єднувати на одному кристалі аналогову, цифрову та обробляючу частини. Крім цього, використання підсилювачів всередині матриці значно збільшує якість та контраст отриманих зображень. Такі матриці набагато дешевші у виробництві та мають кращі характеристики ніж CCD матриці(матриці із зарядовим зв'язком). У сенсорів у тепловізорах використовують болометричні матриці, які складають з теплочутливих елементів. Ці елементи змінюють опір при нагріванні внаслідок поглинання електромагнітної енергії.

Лідар також має схожу будову, чутливим елементом є фотодіод, однак головна відмінність полягає у тому, що камера є пасивною структурою, тобто сприймає світло від інших джерел, а лідар сприймає власний випущений струмінь. На основі явища відбиття світла та його розсіювання у різних середовищ і побудований принцип роботи лідарів.

Датчик світла та дощу має вже всього один світлочутливий елемент і працює за принципом зміни опору в залежності від ступеня освітленості, а використовуючи додаткове випромінювання відбувається процес визначення чи йде дощ. Будується він також на основі явища відбиття світла, а краплі дощу змінюють напрямок поширення світла.

Для позиціювання джерела світла використовується гібридний кроковий двигун. Такий вибір зумовлено досить високою точністю позиціювання та можливістю роботи у мікро-кроковому режимі. Для управління двигуном використовується драйвер крокового двигуна.

Звукова сигналізація реалізується з використанням динаміка, а отже, для роботи, сигнал перетворюється в аналоговий вигляд за допомогою цифро-аналогового перетворювача та підсилюється. Так само управляються і джерело інфрачервоного світла та джерело світла, для підсвічування виявленого об'єкта.

Для роботи обігрівача та соленоїда потрібно комутувати сигнали з більшими значеннями сили струму та напруги, тому для управління ними використовуються реле, куди надходить управляючий сигнал.

Для передачі інформації від камери, тепловізора та лідару необхідно використовувати канали зв'язку з протоколами, що забезпечують високу швидкість передачі даних, це зумовлено об'ємами інформації, що генеруються цими давачами. Тому для обміну інформацією використовується протокол MOSTbus. MOSTbus – це високошвидкісний протокол передачі даних, призначений для мультимедійних мереж, оптимізований для використання у автомобільній промисловості. Послідовна шина MOST використовує топологію шини або кільця та синхронний зв'язок для транспортування даних аудіо, відео, голосових та даних сигналів через оптичне волокно (POF) або кабельний провідник.

Мережа MOST здатна керувати до 64 пристроями при кільцевій топології. Функція Plug and Play дозволяє легко налаштовувати та змінювати пристрої MOST. MOST-мережі також можуть бути налаштовані за топологією уявної зірки, а також іншими.. У критичних, для безпеки системах застосовуються надлишкові подвійні алгоритми шифрування. Можливе використання концентраторів та маршрутизаторів.

У мережі MOST один пристрій є майстром синхронізації. Його роль полягає в постійній підтримці кільця з MOST-кадрами. Преамбула надсилається на початку передачі кадру. Інші пристрої використовують преамбулу для синхронізації. Кодування, що базується на синхронній передачі даних, дозволяє здійснювати постійну після синхронізацію для інших пристроїв.

Для зв'язку між іншими елементами системи використовується шина CANbus. CAN – це стандарт промислової мережі, орієнтований, перш за все, на об'єднання в єдину мережу різних виконавчих пристроїв і датчиків. Режим передачі - послідовний, широкомовний, пакетний.

Безпосередньо стандарт CAN компанії Bosch визначає передачу у відриві від фізичного рівня – він може бути яким завгодно, наприклад, радіоканалом або оптичним волокном. Але на практиці під CAN-мережею зазвичай мається на увазі мережа топології «шина» з фізичним рівнем у вигляді витої пари. Передача ведеться

кадрами, які приймаються всіма вузлами мережі. Для доступу до шини випускаються спеціалізовані мікросхеми - драйвери CAN-шини.

CAN є синхронної шиною з типом доступу Collision Resolving (CR, дозвіл колізії), який, на відміну від Collision Detect (CD, виявлення колізії) мереж (Ethernet), детерміновано (пріоритетно) забезпечує доступ на передачу повідомлення, що особливо важливо для промислових мереж управління. Передача виконується кадрами. Корисна інформація в кадрі складається з ідентифікатора довжиною 11 біт (стандартний формат) або 29 біт (розширений формат) і поля даних довжиною від 0 до 8 байт. Ідентифікатор говорить про вміст пакета і використовується для визначення пріоритету при спробі одночасної передачі декількома мережевими вузлами.

Всі вузли в мережі повинні працювати з однією швидкістю. Стандарт CAN не визначає швидкостей роботи, але більшість як окремих, так і вбудованих в мікроконтролери адаптерів дозволяють плавно змінювати швидкість в діапазоні від 20 Кбіт в секунду до 1 Мбіта на секунду. Існують рішення, що виходять далеко за рамки даного діапазону.

До переваг використання такого протоколу можна віднести можливість роботи у режимі реального часу, просту реалізацію, високу стійкість до завад, арбітраж доступу в мережу без втрати пропускну здатності, широкий діапазон робочих швидкостей та високу популярність та підтримку виробниками.

Недоліками є невеликий об'єм інформації, яку можна передавати в одному пакеті, великий розмір службових даних та відсутність єдиного стандарту.

Для забезпечення зв'язку між системами, що мають високий пріоритет використовується додаткова виділена лінія з протоколом CAN.

7.2 Висновки

У розділі описано функціональні принципи роботи елементів системи. Наведено короткий опис особливостей для кожного з них. Визначено способи передачі даних для різних типів елементів.

8 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ

8.1 Опис ідеї проекту

Метою магістерської дисертації є підвищення безпеки кермування автомобілем за рахунок розроблення системи розпізнавання об'єктів з високими показниками якості.

Зміст ідеї, можливі напрямки застосування та основні вигоди, що може отримати користувач товару представлено у таблиці 8.1

Таблиця 8.1 – Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
	1. Побудова автономних транспортних засобів з використанням даного підходу	Підвищення рівня безпеки на дорогах загального користування. Підвищення комфорту під час керування автомобілем.
	2. Використання підходу для модернізації існуючих систем	Підвищення якості роботи існуючої системи без великих вкладень

Запропонований метод відрізняється від існуючих великою швидкістю роботи та відносно невеликою вартістю.

Було визначено перелік техніко – економічних властивостей та характеристик ідей. Потім проведено аналіз потенційно можливих техніко-економічних властивостей та характеристик ідеї та існуючих аналогів. Для проведення такого аналізу, визначено попереднє коло опонентів, що представлені на ринку. Результати досліджень представлено у таблиці 8.2.

Таблиця 8.2 – Визначення сильних, слабких та нейтральних характеристик ідеї

№ п/п	Техніко- економічні характерис- тики ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтра- льна сторона)	S (сильна сторона)
		Мій проект	NVIDIA DRIVE	NXP BlueBox			
1.	Вартість одиниці	Неможливо оцінити, так як представлен а модель тестувалася з використан ням обчислюва- льних ресурсів комп'ютера	Вартість станови ть приблиз но 15000 доларів	Вартіс- ть стано- вить прибли зно 12000 доларів	При масовому виробниц тві ціна може зрости		Вартість готового продукт у менша за рахунок меншої кількості компонентів
2.	Швидкодія (fps)	51	2000	200	Обме- женість обчислю- вальних ресурсів	Швидко дія може бути підвище на, але зі зростан- ням ціни	
3.	Точність розпізнаван- ня(mAp)	56.3%	82%	78%	Нижча точність		

Продовження таблиці 8.2

№ п/п	Техніко- економічні характерис- тики ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтра- льна сторона)	S (сильна сторона)
		Мій проект	NVIDIA DRIVE	NXP BlueBox			
4.	Давачі інформації	Лідар, камера та тепловізор	Камера	Лідар та камера			Використан- ня поєднання дає більше можливос- тей
5.	Можливість роботи у нічний час	Так, при цьому точність не зменшується	Так, при незначно му зниженні точності	Так, при незначн ому знижен- ні точнос- ті			Тепловізо дає змогу побачити вночі не тільки людей, але і тварин
6.	Можливість роботи у режимі реального часу	Так	Так	Так		Всі методи, повинні відповід ати цьому крите- рію	

Продовження таблиці 8.2

№ п/п	Техніко- економічні характерис- тики ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтра- льна сторона)	S (силь- на сторо- на)
		Мій проект	NVIDIA DRIVE	NXP BlueBox			
7.	Можливі- сть встановле- ння як додатко- вого обладнання	Так	Так, компанія Tesla викорис- товує за замовчу- ванням	Так		Даний тип систем передбачає можли- вість встановлен- ня як опції	
8.	Сертифі- кація	Ні	Так	Так	Для отримання сертифікації необхідно проводити багато часі тестування спочатку на закритих полігонах, а тільки потім на дорогах загального користуван- ня.		

На основі даних, наведених у таблиці, можна зробити висновок, що запропонована ідея є конкурентоспроможною, адже володіє усіма необхідними характеристиками, проте має нижчі показники якості, які можна підвищити використавши платформу з більшою обчислювальною потужністю.

8.2 Технологічний аудит ідеї проекту

У таблиці 8.3 приведено показники технічного аудиту ідеї проекту, та визначено за якою технологією буде виготовлено продукт, проведено аналіз існування необхідних технологій та їх доступність.

Таблиця 8.3 – Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології реалізації	Наявність технологій	Доступність технологій
1.		Використання лідару	Наявна велика кількість різної структури та характеристик	Доступні
2.		Використання радару	Наявні деякі з зразків	Доступні
Обрана технологія базується на використанні лідару				
3.		Написання власного алгоритму обробки з використанням зручних технологій	Потрібно розробляти	Потрібно розробляти
4.		Модифікація існуючих алгоритмів	Потрібно розробляти	Потрібно розробляти

Було обрано технологію модифікації існуючого алгоритму YOLOv3 на мережі DarkNet використовуючи мову програмування Python.

З огляду на дані таблиці проект має змогу технічної реалізації.

8.3 Аналіз ринкових можливостей стартап-проекту

Визначено ринкові можливості, які можна використати під час ринкового впровадження проекту та ринкових загроз, які можуть зашкодити реалізації проекту.

В таблиці 8.4 показано результати аналізу попиту та стану справ на ринку.

Таблиця 8.4 – Попередня характеристика потенційного ринку стартап-проекту

№ п/п	Показники стану ринку (найменування)	Характеристика
1.	Кількість головних гравців, од	12
2.	Загальний обсяг продаж, грн/ум.од	288 000
3.	Динаміка ринку (якісна оцінка)	Зростає
4.	Наявність обмежень для входу (вказати характер обмежень)	Є, проходження сертифікації
5.	Специфічні вимоги до стандартизації та сертифікації	Проходження всіх тестів безпеки, проведення великої кількості годин тестів на полігонах та дорогах
6.	Середня норма рентабельності в галузі (або по ринку), %	78

Роблячи аналіз отриманих результатів, можна сказати, що ринок є доволі привабливим для входження.

Наступним етапом є визначення потенційних груп клієнтів, їх характеристики та формування орієнтовного переліку вимог до товару для кожної групи. Результати досліджень представлено у таблиці 8.5.

Таблиця 8.5 – Характеристика потенційних клієнтів стартап-проекту

№ п/п	Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
1.		Люди, що цікавляться сучасними трендами автономних транспортних засобів	Бажання використовувати нові технології	Приваблива ціна, високі показники якості
2.		Автоконцерни	Збільшення попиту на свою продукцію	Можливість інтеграції у свої автомобілі, висока якість та швидкодія, наявність сертифікатів та відповідність міжнародним стандартам

Далі, проводимо аналіз ринкового середовища та складаємо таблиці факторів, що сприяють ринковому впровадженню проекту, та факторів, що йому перешкоджають. Результати представлено у таблиці 8.6 та 8.7.

Таблиця 8.6 – Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1.	Поява більш продуктивних методів	Можлива поява та стрімкий розвиток нового методу, що матиме кращі характеристики	Спроба підвищити якість системи
2.	Відсутність інвестицій	При відсутності інвестицій неможливий початок виробництва у промислових масштабах	Пошук інших джерел інвестицій
3.	Не проходження сертифікації	За відсутності сертифікатів та невідповідності стандартам неможливо встановлювати такі засоби для автомобілів	Вдосконалення системи та перепроходження сертифікації та стандартизації

Таблиця 8.7 – Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1.	Сертифікація та стандартизація	Вдало пройти сертифікаційну комісію та стандартизація	Можливість виходу на міжнародний ринок

Продовження таблиці 8.7

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
2.	Отримання контракту	Підписання договору на постачання системи з великим автоконцерном	Збільшення виробничих потужностей
3.	Інвестиції	Отримання коштів від інвесторів	Вкладення інвестицій у нові технології

Далі проводимо аналіз пропозиції та визначаємо рівень конкуренції. Результати представлено у таблиці 8.8

Таблиця 8.8 – Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1. Вказати тип конкуренції - чиста	Майже кожна автомобільна компанія розробляє такі рішення, крім того, технологічні компанії пропонують свої уніфіковані рішення	Розвиток технологічності та інвестування у вивчення розробку нових підходів
2. За рівнем конкурентної боротьби - міжнародний	Всі рішення представлені на міжнародному ринку	Великі масштаби виробництва
3. За галузевою ознакою - міжгалузева	Стосується не тільки ринок авто, але і технологічний	Можливість вести бізнес в одній із галузей

Продовження таблиці 8.8.

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
4. Конкуренція за видами товарів - товарно-родова	Створення товарів однієї категорії	Ведення активної маркетингової діяльності
5. За характером конкурентних переваг - нецінова	Споживачі готові платити більше за більш якісну систему	Спрямування лівової частки доходів на розвиток технологій
6. За інтенсивністю - марочна	Система призначена для спільної цільової аудиторії	Введення активної маркетингової діяльності

Аналіз конкуренції в галузі за моделлю 5 сил М. Портера приведено у таблиці 8.9.

Таблиця 8.9 – Аналіз конкуренції в галузі за М. Портером

	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
Складові аналізу	NVIDIA DRIVE, Tesla, NXP BlueBox	Немає	Немає	Немає	Немає
Висновки:	Конкурентна боротьба досить висока, так	Є можливість виходу на ринок після проведення	Ні	Ні	Немає

Продовження таблиці 8.9

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	як головні конкуренти мають добру репутацію та давно на ринку	досліджень та сертифікації			

З огляду на конкуренцію, робота на ринку можлива, проте доведеться конкурувати з гігантами технічної сфери, які довгий час знаходяться на ринку.

Визначимо фактори конкурентоспроможності, їх наведено у таблиці 8.10.

Таблиця 8.10 – Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
1.	Використання технології поєднання дачів інформації	Представлена система використовує одночасне використання кольорової камери та лідару, що значно підвищує точність позиціонування як самого авто, так і об'єктів навколо нього, якісні характеристики роботи алгоритму.
2.	Низька собівартість	Використання лише однієї камери, а також можливість використання твердотілого радару значно знижують витрати на комплектуючі
3.	Можливість роботи в темний час	Використання тепловізора гарантує збереження якісних показників системи на тому самому рівні,

Продовження таблиці 8.10

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
		крім того, використання тепловізора дає можливість більш точної роботи системи при недостатньому освітлені
4.	Можливість роботи в несприятливих погодних умовах	За рахунок використання лідару збільшується здатність протистояти несприятливим погодним умовам, так як якість одержаних моделей знижується не набагато, на відміну від систем, що використовують тільки оптичне сприйняття інформації

За визначеними факторами конкурентоспроможності проводимо аналіз сильних та слабких сторін стартап-проекту. Порівняльний аналіз представлено у таблиці 8.11

Таблиця 8.11 – Порівняльний аналіз сильних та слабких сторін проекту

№ п/ п	Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні з власним проектом						
			-3	-2	-1	0	+1	+2	+3
1.	Використання технології поєднання дачів інформації	18	+						
2.	Низька собівартість	20	+						
3.	Можливість роботи в темний час	17				+			
4.	Можливість роботи в несприятливих погодних умовах	18				+			

Фінальним етапом ринкового аналізу можливостей впровадження проекту є SWOT – аналіз, на основі виділених ринкових загроз та можливостей, та сильних та слабких сторін. SWOT – аналіз представлено у таблиці 8.12.

Таблиця 8.12 – SWOT – аналіз стартап-проекту

Сильні сторони: Низька собівартість Гарні показники якості роботи	Слабкі сторони: Наявність лише моделі системи без фізично реалізації Неможливість пройти сертифікацію
Можливості: Створення фізичної моделі	Загрози: Конкуренція з боку компаній з великим досвідом та авторитетом

На основі SWOT – аналізу розроблено альтернативи ринкової поведінки для виведення стартап – проекту на ринок та орієнтовний час їх ринкової реалізації. Розроблені альтернативи показано у таблиці 8.13

Таблиця 8.13 – Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1.	Реалізація фізичної моделі	Малоймовірно	24 місяці
2.	Вихід на ринок не як готове рішення, а як потенційна модель	Ймовірно	1 місяць
3.	Заключення контракту з автовиробником	Малоймовірно	3 місяці

Аналізуючи наявні перспективи найвигіднішим варіантом є реалізація продукту не у вигляді готового рішення, а у вигляді моделі з реалізованими та протестованими алгоритмами.

8.4 Розроблення ринкової стратегії проекту

На першому кроці визначимо стратегії охоплення ринку. Опис цільових груп представлено у таблиці 8.14

Таблиця 8.14 – Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
1.	Автоконцерни	Низька	25%	Середня	Висока
2.	Власники авто	Середня	45%	Низька	Висока
3.	Продуктові компанії, що займаються готовими рішеннями та тюнінг ательє	Висока	75%	Середня	Середня
4.	Дослідницькі центри у сфері створення	Середня	60%	Низька	Середня

Продовження таблиці 8.14

	систем автономного керування				
Які цільові групи обрано: продуктові компанії, що займаються готовими рішеннями та тюнінг ательє.					

Отже, обрана стратегія концентрованого маркетингу.

Для роботи в обраному сегменті ринку сформулюємо базову стратегію розвитку.

Визначення базової стратегії розвитку представлено в таблиці 8.15

Таблиця 8.15 – Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспромо- жні позиції відповідно до обраної альтернативи	Базова стратегія розвитку*
1.	Вихід на ринок не як готове рішення, а як потенційна модель	Концентрова- ного маркетингу	Можливість виходу на ринок без створення фізичного готового рішення	Стратегія диференціації

У таблиці 8.16 представлено визначення базової стратегії конкурентної поведінки.

Таблиця 8.16 – Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект «першопрохідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки*
1.	Ні	Шукати нових	Ні	Стратегія заняття конкурентної ніші

Визначення стратегії позиціювання представлено у таблиці 8.17

Таблиця 8.17 – Визначення стратегії позиціювання

№ п/ п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспромо жні позиції власного стартап- проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
1.	Можливість використан ня як додаткової опції. Оптимальна ціна із співвідноше	Стратегія диференц іації	Використання додаткових засобів отримання інформації. Низька собівартість.	Низька собівартість. Поєднання технологій отримування інформації. Аналіз об'єднаних даних.

Продовження таблиці 8.17

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувану комплексну позицію власного проекту (три ключових)
	<p>Висока ціна – якість.</p> <p>Високі показники якості та швидкодії.</p> <p>Наявність сертифікатів та відповідність стандартам якості та безпеки</p>			

8.5 Розроблення маркетингової програми стартап – проекту

Сформуємо маркетингову концепцію продукту. Визначення ключових переваг концепції потенційного продукту представлено в таблиці 8.18.

Таблиця 8.18 – Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
1.	Швидкість роботи системи	Система забезпечує швидкість достатню для роботи у режимі реального часу	Потрібно ще більше підвищити швидкість роботи системи, аби вона була вище середньої.
2.	Точність розпізнаван- ня	Завдяки поєднанню сенсорів, система має досить високі показники якості середньої точності	Використання поєднання двох типів сенсорів.
3.	Низька собівартість	Завдяки способу обробки інформації є можливість використання бюджетних засобів отримання та обробки інформації	Більш низька собівартість, зумовлена меншими вимогами до апаратної частини системи.

Продовження таблиці 8.18

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
4.	Можливість встановлення як опції	Система розроблена таким чином, що її можна використовувати як додаткове обладнання	Уніфікація використовуваних технологій та легкість налаштування системи.

Далі розробимо трирівневу маркетингову модель товару. Опис трьох рівнів моделі товару представлено у таблиці 8.19

Таблиця 8.19 – Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Підсистема розпізнавання об'єктів для автомобільних систем автономного управління, що використовує поєднання камери та лідару для підвищення точності та якості розпізнавання.		
	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. Швидкість роботи 54 кадри/сек		
	2. Середня точність – 67 %		
	3. Низька собівартість		
	Якість: показник швидкодії та точності показує, що є можливість використання системи у реальному житті.		
	Пакування		
	Марка: Bird's-eye view		

Продовження таблиці 8.19

Рівні товару	Сутність та складові
III. Товар із підкріпленням	Програмне забезпечення
Потенційний товар буде захищено від копіювання за допомогою використання патентів.	

На наступному кроці визначимо цінові межі, якими необхідно керуватися при встановленні ціни на потенційний товар. Визначення меж встановлення ціни показано у таблиці 8.20.

Таблиця 8.20 – Визначення меж встановлення ціни

№ п/п	Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1.	Відсутні	15 000 – 20 000 \$	4 000 \$/місяць	3 000 – 5 000 \$

Далі. Формування системи збуту – таблиця 8.21.

Таблиця 8.21 – Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1.	Замовлення виготовлення системи, оплата та встановлення	Інформування споживачів	Один рівень	Комбінована

Розробимо концепцію маркетингової комунікації. Її розробку представлено у таблиці 8.22

Таблиця 8.22 – Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
1.	Знайомі з аналогами відомих виробників, таких як Tesla	Інтернет	Інноваційне рішення. Низька собівартість. Поєднання технологій	Показати переваги рішення	Інформаційні відео

8.6 Висновки

Підсумовуючи аналіз стартап – проекту, можна стверджувати, що існує можливість ринкової капіталізації проекту. Ринок автономного транспорту розвивається швидкими темпами, є постійний попит як зі сторони кінцевих споживачів, так і зі сторони виробників, що шукають шляхи покращення своєї продукції. У вигляді готової системи рентабельність є не дуже високою, адже для підготовки такої системи необхідно залучати крупних інвесторів, проводити додаткові дослідження і тести, а потім сертифікувати та стандартизувати систему. Тому, кращим варіантом комерціалізації проекту є вихід на ринок у вигляді програмно-апаратної реалізації моделі. Такий підхід одразу відкидає один із головних бар'єрів входу, про який йшлося вище – це сертифікати міжнародних комітетів з питань безпеки та автономних транспортних засобів.

Конкуренція на досліджуваному ринку чесна, при повній реалізації проекту, отримана система була б конкурентоспроможна. Як одну з головних потенційних груп клієнтів слід розглядати тюнингові ательє, що займаються встановленням додатковим систем.

Подальша імплементація проекту є доцільною з точки зору розвитку галузі загалом, а також з точки зору створення тестової фізично реалізованої моделі для дослідження роботи моделі у реальному житті.

ВИСНОВКИ

Магістерська дисертація присвячена задачі підвищення рівня безпеки на дорогах та рівня комфорту при керуванні автомобілем за рахунок розроблення підсистеми розпізнавання об'єктів системи для автомобільних систем автономного управління, яка використовує алгоритми, що мають кращі показники якості та швидкості роботи.

Для ознайомлення з принципами побудови систем розпізнавання об'єктів було зроблено огляд існуючих методів та алгоритмів, що використовуються на сьогоднішній день. Крім цього, було розглянуто існуючі системи, які використовують провідні концерни автомобілебудування. Огляд існуючих рішень представлено у розділі 2. На основі аналізу наведеної інформації було прийнято рішення про структуру системи, що передбачає одночасне використання кольорової камери та лідару.

Вибір використовуваного алгоритму здійснювався на основі аналізу результатів моделювання цих алгоритмів. Опис алгоритмів та їх математична модель представлена у розділі 3. Для порівняння алгоритмів було реалізовано п'ять найбільш розповсюджених. Для моделювання було використано середовище MATLAB, тренування та тестування мереж відбувалося на основі датасету Caltech Pedestrian Dataset, який є одним із найбільших, однак містить тільки анотацію для виявлення пішоходів. Результати моделювання, а також отримані результати та їх порівняльний аналіз наведено у розділі 4. Опираючись на отримані дані, було прийнято рішення про використання алгоритму YOLO.

Апробація результатів порівняння алгоритмів та вибору оптимальної моделі впроваджено на VII міжнародній науково – практичній конференції Winter InfoCom Advanced Solutions 2019.

Після вибору структури системи та алгоритму було реалізовано програмну модель системи. Більш детально про принципи роботи та побудови моделі можна переглянути у розділі 5. Також, було проведено моделювання роботи системи та отримані результати ефективності – швидкодія системи 50.4 кадрів/сек, а середня

точність становить 56.3 %. Такі показники забезпечують можливість роботи системи у режимі реального часу та використання моделі у реальних системах. За рахунок використання лідару були отримані високі показники точності обчислення позиціювання як самого авто, так і об'єктів, що знаходяться навколо нього.

Для повного опису системи було розроблено та описано схему електричну структурну та схему електричну функціональну з позиції побудови повної системи автономного управління. Розроблення та опис схем наведено у розділах 6 та 7.

Було проведено дослідження можливості комерціалізації запропонованої ідеї. Для цього, проведено дослідження стартап-проекту за різними показниками, оцінено можливість виходу на ринок, стан та тип конкуренції, можливі перешкоди та виділено переваги нашого продукту відносно конкурентів. На основі аналізу було зроблено висновок про доцільність подальшої імплементації проекту. Розроблення стартап – проекту описано у розділі 8.

Отже, запропонована модель системи розпізнавання об'єктів є досить ефективною в порівнянні з аналогічними системами та має можливість бути реалізованою фізично. Отримані результати підтверджують правильність обраного методу та алгоритму.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Статистика дорожньо-транспортних пригод [Електронний ресурс]: Режим доступу: <http://www.autolivnightvision.com/statistics/> - 12.02.2018 р.
2. Hamer, M. (27 August 2005). "Stopping the slaughter of innocent pedestrians". New Scientist (2514).
3. Офіційний сайт компанії Volvo [Електронний ресурс]: Режим доступу: <https://www.volvocars.com/rs/support/manuals/s60/2019/safety/safety/pedestrian-protection-system> - 12.08.2019 р.
4. Broggi et al. Stereo-based preprocessing for human shape localization in unstructured environments // IEEE Intell. Veh. Symp. – Ohio, USA. – 2003. – P. 410 – 415.
5. Jabri S., Duric Z., Wechsel H., Rosenfeld A. Detection and localisation of people in video images using adaptive fusion of color and edge information // IEEE Int. Conf. Patt. Rec., Fairfax. – 2000. – P. 627 – 630.
6. Fengliang Xu, Kikuo Fujimura Pedestrian Detection and Tracking with Night Vision // Intelligent Transportation Systems, IEEE Transactions. – 2005. – Vol. 6, Issue 1. – P. 63 – 71.
7. Wang, X.Z.; Li, J.; Li, H.J.; Shang, B.X. Obstacle detection based on 3d laser scanner and range image for intelligent vehicle. J. Jilin Univ. (Eng. Technol. Ed.) 2016, 46, 360–365.
8. Xue, J.R.; Wang, D.; Du, S.Y. A vision-centered multi-sensor fusing approach to self-localization and obstacle perception for robotic cars. Front. Inf. Technol. Electron. Eng. 2017, 18, 122–138.
9. Офіційний сайт медіа відділу Daimle Chrysler [Електронний ресурс]: Режим доступу: <https://media.daimler.com/marsMediaSite/en/instance/ko/Assistance-systems-help-to-ensure-safety-Now-also-in-the-Econic-Active-Brake-Assist-4--emergency-brake-assist-system-with-pedestrian-detection.xhtml?oid=40336482> - 18.08.2019 р.
10. Офіційний сайт MOBILEYE VISION TECHNOLOGIES LTD [Електронний ресурс]: Режим доступу: <https://www.israelbizreg.com/mobileye-vision-technologies-ltd> - 18.08.2019 р.

11. Chad Carson, Megan Thomas, Serge Belongie, Joseph M. Hellerstein, and Jitendra Malik. Blobworld: A System for Region-Based Image Indexing and Retrieval.
12. Nello Cristianini, John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. — Cambridge University Press, 2000. — ISBN 978-1-139-64363-4.
13. Офіційний сайт Mercedes – Benz [Електронний ресурс]: Режим доступу: <https://www.mercedes-benz.com/en/innovation/protection-in-accident-scenarios-mercedes-benz-pre-safe/> - 18.08.2019 р.
14. Офіційний сайт Toyota [Електронний ресурс]: Режим доступу: <https://www.toyota.com.au/toyota-safety-sense/pre-collision-safety-system> - 18.08.2019
15. Night vision assist [Електронний ресурс]: Режим доступу: <https://www.audi-mediacentr.com/en/photos/night-vision-assist-52988> - 18.08.2019 р.
16. BMW Night Vision Technology [Електронний ресурс]: Режим доступу: <https://www.edmontonbmw.com/bmw-night-vision-in-edmonton/> - 18.08.2019 р.
17. Subaru EyeSight [Електронний ресурс]: Режим доступу: <https://www.subaru.com/engineering/eyesight.html> - 18.08.2019 р.
18. Golson, Jordan; Bohn, Dieter (2016-10-19). "All new Tesla cars now have hardware for 'full self-driving capabilities'". The Verge. Retrieved 2016-10-22.
19. "Enhancing Safety and Convenience with Summon". Tesla Motors. 2016-02-08. Retrieved 2016-04-18.
20. Cutress, Dr Ian. "Hot Chips 31 Live Blogs: Tesla Solution for Full Self Driving". www.anandtech.com. Retrieved 2019-08-21.
21. Офіційний сайт IIHS [Електронний ресурс]: Режим доступу: <https://www.iihs.org/news/detail/iihs-prepares-to-test-pedestrian-detection-systems> - 25.08.2019 р.
22. Ron Brinkmann (1999). The Art and Science of Digital Compositing. Morgan Kaufmann. p. 184. ISBN 978-0-12-133960-9.
23. L. Bourdev and J. Brandt. Robust object detection via soft cascade. In CVPR, 2005

24. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005.
25. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in ACM MM, 2014
26. Dominguez-Sanchez, A.; Cazorla, M.; Orts-Escolano, S. Pedestrian Movement Direction Recognition Using Convolutional Neural Networks. *IEEE Trans. Intell. Transp. Syst.* 2017, 18, 3540–3548, doi:10.1109/TITS.2017.2726140.
27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, doi:10.1109/TPAMI.2016.2577031.
28. Ordóñez, F.J.; Roggen, D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors* 2016, 16, 115, doi:10.3390/s16010115.
29. L. Wang, H. Lu, X. Ruan, and M.-H. Yang, “Deep networks for saliency detection via local estimation and global search,” in CVPR, 2015.
30. P. Druzhkov and V. Kustikova, “A survey of deep learning methods and software tools for image classification and object detection,” *Pattern Recognition and Image Anal.*, vol. 26, no. 1, p. 9, 2016.
31. Caltech Pedestrian Dataset [Электронный ресурс]: Режим доступа: http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/ - 02.09.2019 г.
32. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, 24–27 June 2014; pp. 580–587.
33. Ren, J.; Chen, X.; Liu, J.; Sun, W.; Pang, J.; Yan, Q.; Tai, Y.W.; Xu, L. Accurate Single Stage Detector Using Recurrent Rolling Convolution. In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017.

34. Xiang, Y.; Choi, W.; Lin, Y.; Savarese, S. Subcategory-aware Convolutional Neural Networks for Object Proposals and Detection. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 24–31 March 2017.
35. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 24–27 June 2014; pp. 580–587.
36. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional Architecture for Fast Feature Embedding. ACM 2014, 675–678, doi:10.1145/2647868.2654889. Steger C., Ulrich M., Wiedemann C. Machine vision algorithms and applications. — Wiley-VCH, 2008. — 360 p. — ISBN 9783527407347.
37. C. L. Zitnick and P. Dollar, “Edge boxes: Locating object proposals from edges,” in ECCV, 2014.
38. R. Girshick, “Fast r-cnn,” in ICCV, 2015.
39. S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards realtime object detection with region proposal networks,” in NIPS, 2015, pp. 91–99.
40. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in CVPR, 2016.
41. Geiger, A.: Are we ready for autonomous driving? the kitti vision benchmark suite. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). CVPR '12, Washington, DC, USA, IEEE Computer Society (2012) 3354–3361
42. Glennie, C.; Lichti, D.D. Static calibration and analysis of the Velodyne HDL-64E S2 for high accuracy mobile scanning. Remote Sens. 2010, 2, 1610–1624.
43. KITTI Vision Benchmark [Электронный ресурс]: Режим доступа: <http://www.cvlibs.net/datasets/kitti/> - 25.09.2019 p.
44. Qi, C.R., Liu, W., Wu, C., Su, H., Guibas, L.J.: Frustum pointnets for 3d object detection from RGB-D data. CoRR abs/1711.08488 (2017)

45. Redmon, J., Divvala, S.K., Girshick, R.B., Farhadi, A.: You only look once: Unified, real-time object detection. CoRR abs/1506.02640 (2015)
46. Beyer, L., Hermans, A., Leibe, B.: Biternion nets: Continuous head pose regression from discrete training labels. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9358 (2015) 157–168

Додаток А – Публікація

CONFERENCE PROCEEDINGS

МАТЕРІАЛИ КОНФЕРЕНЦІЇ


 winter
infoCom 2019

VIII МІЖНАРОДНА НАУКОВО-ПРАКТИЧНА КОНФЕРЕНЦІЯ
З ІНФОРМАЦІЙНИХ СИСТЕМ ТА ТЕХНОЛОГІЙ

Winter InfoCom

Advanced Solutions 2019

8th INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE
ON INFORMATION SYSTEMS AND TECHNOLOGIES

2-3 грудня 2019 року
Україна, Київ

2-3 December 2019
Ukraine, Kyiv

Communications
Cyber Intelligence
Cyber Security

Pe Pwne Express	Sc Silent Circle	Pap Payfone	App AppDome	Atk AttackIQ	SS Security Scorecard
Lo Lookout	Zi Zimperium	Trs Trustlook	Sl Sirin Labs	Ce Cyence	Bd Bay Dynamics
Op OpenPeak	Hy Hypori	As Avast Software	Rn RedSeal Networks	Bt BlitSight Technologies	Ks Kenna Security
Ap Appthority	Apk Appknox	Mm Mobi Magic	Sa SafeBreach	Pn Prevalent Networks	Co Corax

Ah Andreessen Horowitz	Lvp Lightspeed Venture Partners	Kpc Kleiner Perkins Caufield & Byers	Nvp Norwest Venture Partners	Gv Google Ventures	Sca Sequoia Capital
Bn Barracuda Networks	Trst Trusteer	Atg AVG Technologies	Kd Kruze Digital	Moi MobileIron	Ops OpenDNS

On OneLogin	Tse Thycotic Software	Ta Tanium	Bii BitDefender	Zs Zscaler	Sty Shift Technology
Cy Centrify	Nnl Nok Nok Labs	Lor LogRhythm	Cs Code42 Software	Th ThreatMatrix	Ga Guardian Analytics
Be BeyondTrust	Tra TruSona	Cr CrowdStrike	Sn SnoopWall	Dn Distil Networks	Fot Forter
Seu SecureAuth	Iw iWelcome	Dg Digital Guardian	St StackPath	Ko Kount	Ze ZeroFOX
So Secure	Ve Veridu	Av Avesta	Loo LogicMonitor	Ju Jumio	Ri RippleShot
Trll Trulioo	Moq MoQom	Coa CounterTack	Cl Cloudflare	Fe Feedzai	Ra Ravelin
Tn Tempered Networks	Sy Simplified	Cb Carbon Black	Il Illumio	Wo White Ops	Sii Simility

ISBN 978-966-2344-67-7

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ІНСТИТУТ МОДЕРНІЗАЦІЇ ЗМІСТУ ОСВІТИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ
УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ
СІКОРСЬКОГО»**

WINTER INFOCOM ADVANCED SOLUTIONS 2019

МАТЕРІАЛИ

**VIII МІЖНАРОДНОЇ НАУКОВО-ПРАКТИЧНОЇ
КОНФЕРЕНЦІЇ**

З ІНФОРМАЦІЙНИХ СИСТЕМ ТА ТЕХНОЛОГІЙ

CONFERENCE PROCEEDINGS

8th SCIENTIFIC AND PRACTICAL CONFERENCE

КИЇВ, УКРАЇНА

2-3 грудня 2019 року

ПРОГРАМА/PROGRAM

2 грудня	Інформаційні системи та технології	
	Хлівненко М. Писаренко А.	Підсистема розпізнавання об'єктів для автомобільних систем автономного керування
	Зубрицький А.	Проектування архітектури системи дослідження тексту
	Покровський Є. Савчук О. Моргаль О. Похиленко О.	Про моделювання надійності та оцінювання в системі хмарних сервісів
	Bodak B. Doroshenko A.	The impact and unforeseen challenges of E-procurement systems in Canada
	Оброблення інформації у складних системах	
	Холодович К. Букасов М.	Автоматизація пошуку помилок у сирих даних та створення SDTM датасетів для медичних досліджень
	Писаренко О. Дорошенко А.	Аналіз коментарів за допомогою машинного навчання
	Poltorak V.	Analysis of the calculus basis boundary for redundant codes
	Інформаційні системи та технології	
3 грудня	Kharabet R. Pysarenko A.	The use of radio-frequency identification in information systems
	Дяченко К. Писаренко А.	Інтегрована інформаційна система моніторингу та керування на основі інтернету речей для розумної ферми
	Теленик А.	Автоматизована система віддаленої інсталяції програмного забезпечення
	Alhawawsha M. Anisimov A.	Developing of the E-government System based on Java for Online Voting
	Безпека та захист інформації	
	Романчук С.	Аналіз особливостей державних стандартів ЕЦП на властивостях еліптичних кривих
	Калитюк Н.	Автентифікація зображень на основі методів цифрового підпису

Підсистема розпізнавання об'єктів для автомобільних систем автономного керування

Хлівненко Михайло
КПІ ім. Ігоря Сікорського
Київ, Україна
xlivnenko.michael@gmail.com

Писаренко Андрій
КПІ ім. Ігоря Сікорського
Київ, Україна
andrew.pisarenko@gmail.com

Анотація. Наведено аналіз існуючих апаратних методів розпізнавання об'єктів для систем автономного керування. Представлено дослідження показників якості алгоритмів розпізнавання машинного навчання. На основі аналізу запропоновано структуру моделі системи розпізнавання об'єктів для автомобільних систем автономного керування.

Ключові слова: системи автономного керування, алгоритми розпізнавання об'єктів, ACF, YOLO, лідар.

ВСТУП

Останнім часом, зростає кількість дорожньо-транспортних пригод пов'язаних з пішоходами та за таких обставин, що водій не має змоги вчасно зреагувати на виникнення небезпечних ситуацій [1]. Світові гіганти автомобілебудування приділяють багато уваги безпеці руху, кожного року представляють нові системи, вдосконалюють існуючі. Дані системи аналізують навколишній простір та у разі виникнення небезпечних ситуацій сповіщають про це водія, або самостійно втрачаються у процес керування.

З огляду на такий стан речей, для розроблення систем розпізнавання об'єктів використовують сучасні апаратні рішення, що здатні забезпечити усі необхідні умови. Проте, не менш важливою частиною систем є програмне забезпечення. Кожного року представляють нові алгоритми, методи та підходи до машинного навчання, кожен з яких, має свої переваги та недоліки. Проблема, що виникає перед розробниками – це необхідність вибору певного підходу та алгоритму, що забезпечив би при цьому усі висунуті вимоги, при цьому, за час створення, відлагодження, тестування та сертифікації системи такі підходи зазвичай втрачають актуальність.

Отже, постає питання порівняння сучасних методів розпізнавання та на основі аналізу їх показників створення такого підходу, який би був легко інтегрованим, при цьому, мав високі показники якості роботи та мінімальний час навчання.

ОГЛЯД ІСНУЮЧИХ ПІДХОДІВ ТА МЕТОДІВ

Вирішуючи проблему виявлення об'єкта та його місцезнаходження, провідні виробники прийшли до розширення можливостей систем за рахунок використання різних апаратних засобів, що значно

збільшують кількість необхідної інформації, проте, одразу ж виникає проблема з обробкою такого об'єму даних.

На сьогодні, існує три базових апаратних засоби, що використовуються в автомобільних системах – камери (включаючи тепловізори), лідари та радары.

Камери є універсальним засобом отримання інформації з порівняно невисокою ціною та високою ефективністю. До переваг камер можна віднести невеликий обсяг генерованої інформації та відносно невеликі вимоги до потужності апаратної платформи. Також, камери мають здатність розрізнення кольорів, що позитивно впливає на якість розпізнавання. Досить часто, виробники використовують камери, що мають теплову роздільну здатність – тепловізори. Такий підхід дозволяє ефективно орієнтуватися у темний час доби.

Лідар – це світовий радар, що дає змогу отримувати об'ємну модель простору навколо транспортного засобу. Крім цього, лідар дає змогу оцінювати відстань до об'єктів. Основною проблемою є великий об'єм генерованої інформації, що вимагає великої апаратної потужності для подальшої обробки, а також порівняно висока вартість. На роботу лідару можуть впливати погодні умови, наприклад, дощ та туман [2].

Радар – це сенсорна система, яка використовує радіохвилі для визначення швидкості, дальності та кута об'єкта. Радар використовує невеликий об'єм даних та потребує менше обчислювальної потужності ніж камера. За своєю точністю він поступається лідару, проте, має можливість працювати у несприятливих погодних умовах.

ОГЛЯД АЛГОРИТМІВ РОЗПІЗНАВАННЯ ТА ЇХ АНАЛІЗ

Для розпізнавання об'єктів застосовуються будь-які з доступних класифікаторів. Зазвичай, виробники використовують потужні алгоритми попередньої обробки та будь-який з доступних класифікаторів. Часто в якості класифікатора використовують нейронні мережі. В більшості розглянутих систем – це Support Vector Machine (SVM) [3], причому ні в одній роботі немає аргументації, чому слід застосовувати саме цей тип нейронних мереж. Порівняльний аналіз різних класифікаторів також відсутній.

Однак, такий підхід не є досконалим, тому постає

проблема створення універсального засобу для розпізнавання.

Для аналізу алгоритмів, було реалізовано програмні моделі детекторів за допомогою MATLAB. Для аналізу було обрано наступні алгоритми: ACF, R-CNN [4], Fast-RCNN [5], Faster-RCNN [6] та YOLO [7]. Такий вибір зумовлений можливістю порівняти двоетапні (використовує певний фрагмент зображення) та одноетапні детектори, а також легкістю їх реалізації та відносно високою ефективністю.

Для тренування було обрано найбільший з сьогодні існуючих датасетів – Caltech Pedestrian [8], що містить понад 10 годин реальних відео.

Експерименти використовують наступну процедуру оцінювання. Трекери ініціалізуються в першому кадрі відеопослідовності і відслідковують об'єкт, що нас цікавить до кінця. Потім отриману траєкторію порівнюють із еталонною використовуючи такі параметри: тривалість треку та точність. У таблиці 1 представлено показники якості для тренуваних моделей з використанням різних алгоритмів.

Таблиця 1

Алгоритм	Показники якості	
	Тривалість треку	Точність
ACF	0.96	0.88
R-CNN	0.95	0.15
Fast R-CNN	1	0.29
Faster R-CNN	1	0.33
YOLO	1	0.77

Також, було оцінено швидкість обробки кадрів, результати показано на рис. 1.

Аналізуючи отримані результати, можна сказати, що кращим варіантом для використання в системах

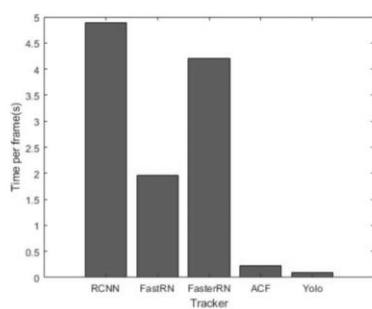


Рис. 1. Швидкість обробки у кадрах/секунду

розпізнавання реального часу є алгоритм YOLO (You only look once), адже він має найвищу швидкість роботи та досить високу точність.

СТРУКТУРА СИСТЕМИ

З огляду на представлену інформацію, найкращим варіантом побудови системи є використання камери з лідаром. Такий підхід

забезпечить високу ефективність розпізнавання об'єктів та дасть змогу побудувати об'ємну модель навколишнього простору, яка робить можливим більш точно визначати координати об'єктів, їх положення та швидкість відносно автомобіля. У якості алгоритму для розпізнавання об'єктів на зображеннях, отриманих з камери використовується YOLO.

ВИСНОВКИ

На основі аналізу існуючих методів та алгоритмів розпізнавання пішоходів було запропоновано модель системи розпізнавання об'єктів для автомобільних систем автономного керування в основі якої поєднано використання камери та лідару. У якості алгоритму розпізнавання використовується YOLO. Для передбачення траєкторії руху об'єкта використовується фільтр Калмана.

ЛІТЕРАТУРА

1. Страховий інститут безпеки дорожнього руху [Електронний ресурс] // IIHS. – 2019. – Режим доступу до ресурсу: <https://www.iihs.org>.
2. Volz, B.; Behrendt, K.; Mielenz, H.; Gilitschenski, I.; Siegart, R.; Nieto, J. A data-driven approach for pedestrian intention estimation. In Proceedings of the International Conference on Intelligent Transportation Systems. IEEE, Rio de Janeiro, Brazil, 1–4 November 2016; pp. 2607–2612, doi:10.1109/ITSC.2016.7795975.
3. Sun, W.; Zhu, S.; Ju, X.; Wang, D. Deep learning based pedestrian detection. In Proceedings of the Chinese Control And Decision Conference (CCDC), Shenyang, China, 9–11 June 2018; pp. 1007–1011.
4. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 24–27 June 2014; pp. 580–587.
5. Li, X.; Li, L.; Flohr, F.; Wang, J.; Xiong, H.; Bernhard, M.; Pan, S.; Gavrila, D.M.; Li, K. A unified framework for concurrent pedestrian and cyclist detection. IEEE Trans. Intell. Transp. Syst. 2017, 18, 269–281, doi:10.1109/TITS.2016.2567418.
6. Brazil, G.; Yin, X.; Liu, X. Illuminating Pedestrians via Simultaneous Detection and Segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4960–4969, doi:10.1109/ICCV.2017.530.
7. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. European Conference on Computer Vision; Springer: Cham, Switzerland, 2016; pp. 21–37, doi:10.1007/978-3-319-46448-0_2.
8. Caltech Pedestrian Detection Benchmark [Електронний ресурс] // Caltech. – 2019. – Режим доступу до ресурсу: http://www.vision.caltech.edu/Image_Datasets/Caltech_Pedestrians/